The Dissertation Committee for Thaddeus Bradley Czuba Certifies that this is the approved version of the following dissertation:

Binocular mechanisms underlying the processing of three-dimensional visual motion.

Committee:

Lawrence K. Cormack, Co-Supervisor

Alexander C. Huk, Co-Supervisor

Wilson S. Geisler

Jonathan W. Pillow

Nicholas J. Priebe

Binocular mechanisms underlying the processing of three-dimensional visual motion.

by

Thaddeus Bradley Czuba, B.S. Psy.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2012

Acknowledgements

This dissertation would not have been possible without the collaborative effort of my supervisors, Larry Cormack & Alex Huk, and our longtime co-author Bas Rokers. I would like to thank Alex & Larry for fostering such an amazing academic environment throughout my graduate career. From entertaining my desire to run psychophysical experiments on a big screen television, to being able to laugh off the fact that a first year graduate student missed his plane to the conference (sorry!), I don't think I can imagine a more supportive set of advisors. You have both set amazing examples of how to live life —within and beyond the university setting—and I look forward to our paths crossing in the years to come.

Thank you to the faculty and staff of the UT Center for Perceptual Systems. It has been an honor to work with such a great group of people in such a unique and collaborative setting. A big thanks to the members of my dissertation committee: Bill Geisler, Jonathan Pillow, & Nicholas Priebe for helping to shape this dissertation into the document it is today. Thank you to Jeff Luci and the staff of the Imaging Research Center for your tireless efforts to develop and maintain our neuroimaging equipment & facilities. I would also like to thank professors Jason Gold (Indiana University) and Don MacLeod (UCSD) for their tutelage in the early years of my scientific career; without your encouragement and support I would not be where I am today.

I've had the fortune to work with many excellent graduate students and postdocs that have provided intellectual stimulation and extracurricular adventure throughout the years: Bas Rokers, for research inspiration and Socratic guidance during the first half of my graduate career; Johannes Burge, for setting a great example of how to strike a balance in postdoctoral life; lab members Miriam Meister, Jacob Yates, & Leor Katz, for providing helpful feedback on countless poster/talk drafts and stimulating conversation over the years. A special thanks to my first lab mates, Dirk Beer & Erin Krizay, for their team building efforts, and Erin Shubel, for being my serendipitous scientific coordinator

over the years. I certainly wouldn't have made it through graduate school without a steady regimen of disc golf and trivia competitions with an unbeatable group of friends: Matt Brooks, Conner Bryan, Michael Buhrmester, Amelia Hall, Dean Kirson, Jackson Lang, Scott Liening, Brian McCann, Jenni Pacheco, Brian Sullivan, Kyle Walsh, Randy Way, ...and many more. Thank you all for so many great times over the years.

Finally, I'd like to thank my family for decades of encouragement and inspiration. I really couldn't have asked for a more loving upbringing. The tireless support of my parents allowed me to discover and pursue a career in science that wouldn't have been possible without the intellectual curiosity and (occasionally stubborn) independence that they instilled in me from a very young age. Thank you to my sister, brother-in-law, and newborn niece for the inspiration to chase my dreams wherever and as far as they may go. Last but not least, a special thank you to Colleen Cappellini. Despite my disappearing into (and at times sleeping in) the lab for weeks at a time, you stuck by me throughout the many ups and downs of graduate school. Thank you for always being there to remind me that life is as much about what you've done, as it is about the adventures you have along the way.

Binocular mechanisms underlying the processing of three-dimensional visual motion.

Thaddeus Bradley Czuba, Ph.D.
The University of Texas at Austin, 2012

Supervisor: Lawrence K. Cormack & Alexander C. Huk

In this dissertation, I examine binocular 3D motion processing through a series of psychophysical and neuroimaging experiments aimed at uncovering the neural computations involved and their interaction with the known hierarchy of visual motion processing. Two primary binocular cues could be used to compute 3D motion: one based on changing disparities over time (CD), the other based on interocular velocity differences (IOVD). Under normal viewing conditions, both cues coexist and (potentially) provide the same 3D direction information, yet whether CD, IOVD, or both mechanisms exist has distinct implications for how 3D motion is processed along the visual stream.

First, I measured 3D direction discrimination sensitivity is measured for isolated binocular cues under a range of 3D motion speeds and visual eccentricities. Comparison of isolated-cue sensitivity to corresponding combined cue sensitivity (i.e. concurrent IOVD & CD cue stimuli) provided an estimate of relative cue contributions under normal viewing conditions. Second, I conducted a series of motion adaptation experiments to differentiate the neural representation of 2D and 3D directions of motion, and examine the degree to which IOVD or CD mechanisms can account for 3D motion adaptation. Third, I examined the neural locus of 3D motion processing by measuring 3D direction-selectivity throughout a range of visual cortical areas using functional neuroimaging in an event-related paradigm that parallels psychophysical adaptation experiments. Finally, I

discuss the broader implications for the neural mechanisms of binocular 3D motion processing and future experimental directions.

Together, these results reveal that: (1) the IOVD cue is the dominant cue to 3D motion processing across the majority of natural speeds & eccentricities, (2) neural tuning for 3D motion is distinct from 2D motion and can be fully explained by an IOVD mechanism, and (3) the IOVD cue is computed relatively late in the visual processing stream, in areas MT & MST— cortical areas primarily associated with 2D/retinal motion and thought to be beyond the point of binocular combination. The significance of IOVD —but not CD—cues to 3D motion perception motivates a drastic modification to canonical models of motion processing to include the late-stage comparison of eye-specific motion signals.

Table of Contents

List of Figures xi

Chapter 1: Introduction 1

Binocular 3D motion cues 1

Experimental aims 3

Chapter 2: Speed & eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion 6

Introduction 6

Methods & Materials 10

Observers 10

General procedure 10

General stimuli 11

Manipulations of eccentricity and speed 14

Motion cue conditions: FULL, IOVD, and CD 15

Manipulation of 3D motion coherence 18

Experimental design 20

Apparatus and display 20

Data analysis 22

Results 23

Effects of eccentricity at different speeds 24

Effects of speed at different eccentricities 26

Speed by eccentricity (SxE) sensitivity surface 29

Discussion 37

Distinguishing the contributions of the CD and IOVD cues 37

Effects of speed on the CD and IOVD mechanisms 44

Eccentricity effects on the CD and IOVD mechanisms 45

Relation to past work 46

Implications for future work 51

Chapter 3: Neural representation and directionally-selective mechanisms of binocular 3D motion processing 53
Introduction 53
Methods & Materials 56
Observers 56
General procedure 56
General stimuli 57
Manipulation of 3D motion coherence 59
Apparatus and displays 62
Data analysis 64
Experiment 1: Methods 65
2D vs. 3D MAE. 65
Dissociation of 3D MAE from inherited, monocular 2D MAEs 68
Experiment 1: Results 68
Existence of a 3D MAE and comparison to 2D MAE 68
Dissociation of 3D MAE from inherited, monocular 2D MAEs 74
Experiment 2: Methods 78
Binocular cue contribution 78
Isolated cue conditions: CD, IOVD, and 3D-planar. 78
Experiment 2: Results 85
3D MAEs from isolated binocular cues 85
General Discussion 89
Quantitative comparison of all 3D and 2D MAEs 89
Relation to past work 95
Conclusions 96

Chapter 4: To CD or not to CD: Is there a 3D motion aftereffect based on changing disparities? 98

Chapter 5: Neural locus of 3D motion processing 103
Introduction 103
Methods 105

Subjects 105
Functional imaging & processing 105
Visual stimuli 106
Display apparatus 110
fMRI data analysis 110
Preliminary Results 113
Chapter 6: General Discussion 119
Relation to past work 119
Electrophysiology 122
Constraints on putative CD mechanisms 123
Constraints on putative IOVD mechanisms 125
Conclusions & Future Directions 134
Appendix A: Motion processing with two eyes in three dimensions 136
Introduction 136
Methods 138
Observers 138
Apparatus 138
Stimuli 139
Procedure and task 142
Data analysis 143
Results 145
Baseline Experiment: Dichoptic Gratings 146
Dichoptic type I plaid experiment 149
Dichoptic type II plaid experiment 153
Dichoptic pseudoplaid experiments 154
Eye-of-origin & monocular motion control experiments 160
Generalization of effects to naive observers and to a different task 162
Discussion 170
References 176

List of Figures

2.1: Schematic of binocular 3D motion cues 8

2.2: Experimental stimuli & manipulations 13

2.3: Isolation of binocular cues 16

2.4: Effects of eccentricity at different speeds 25

2.5: Effects of speed at different eccentricities 28

2.6: Spatiotemporal sensitivity surfaces 31

2.7: Correlation of isolated and combined cue sensitivity 36

2.8: Confirmation of IOVD cue isolation 41

3.1: Screen capture of the basic stimulus 58

3.2: Motion coherence manipulation (Movie 1) 61

3.3: 3D motion adaptation & test stimulus (Movie 2) 66

3.4: Frontoparallel adaptation & test stimulus (Movie 3) 67

3.5: 3D motion adaptation 70

3.6: Frontoparallel motion adaptation 73

3.7: Monocular motion adaptation 75

3.8: 3D–mono & interocular transfer results 77

3.9: CD-isolating adaptation & test stimulus (Movie 4) 79

3.10: IOVD-isolating adaptation & test stimulus (Movie 5) 81

3.11: 3D-planar adaptation & test stimulus (Movie 6) 84

3.12: Isolated-cue adaptation results 86

3.13: MAE magnitudes across observers & conditions 88

3.14: Psychometric parameter distributions across all conditions 91

4.1: MAE following unidirectional or bidirectional adaptation 100

5.1: Schematic of event-related trial sequence 108

5.2: Example functional time series & blank subtraction 112

5.3: Combined cue adaptation across visual ROIs 115

5.4: 3D direction selectivity index 117
6.1: Binocular plaid (type I) schematic & results 127
6.2: Pseudoplaid stimulus movie 129
6.3: Pseudoplaid stimulus schematic & results 130
6.4: Projective geometry of rigid 3D motion 133
A.1: Dichoptic grating stimulus schematic & results 147
A.2: Dichoptic Type-I & Type-II stimulus schematics & results 151
A.3: Two- and multi-component pseudoplaid stimulus schematics & results 157
A.4: Eye-of-origin & monocular motion control experiments 161
A.5: Naive observer dichoptic grating results & ROC analysis 163
A.6: Naive observer 2AFC results for plaids and pseudoplaids 166
A.7: Demonstration of the pseudoplaid stimulus (Movie 7) 169

Chapter 1: Introduction

Although motion perception has long been a topic of vision research, much of the work has been conducted on fronto-parallel 2D motion (e.g. as presented on a flat computer display). In the real world, motion is rarely constrained to a two-dimensional plane, and often information most relevant to successfully navigating/interacting in real world environments contains substantial components in the third dimension: depth. Owning to its ecological importance, numerous mechanisms exist to reconstruct such three-dimensional information. Specifically, an object's motion through depth can be perceived based on two binocular cues: one disparity-based, and the other velocity-based. While these two cues generally encode identical motion information, they are computationally distinct and have significant implications on how monocular signals are integrated into a single, cyclopean, representation of the visual field. The goal of this dissertation is to examine the contribution of binocular cues to 3D motion processing, with specific aims of understanding (1) what binocular computations support 3D motion processing, (2) how 3D motion is represented by the visual system, and (3) where the neural locus of binocular motion mechanisms exist within the hierarchy of visual motion processing. To do so, I conducted a series of behavioral psychophysics and functional neuroimaging studies using stimuli designed to isolate binocular 3D motion cues.

BINOCULAR 3D MOTION CUES

By taking advantage of slight differences in perspective created by the horizontal offset of the two eyes, the visual system is able to reconstruct a three-dimensional representation of the world around us. Our understanding of such binocular computations has been primarily focused on positional stereopsis, in which an object's position-in-depth can be determined from minute offsets in the position of retinal images projected in the two eyes; known as retinal disparity (Wheatstone, 1838). Expanding on the notion that 2D motion can be computed directly by tracking an object's *retinal position* over time (e.g. Braddick, 1974), binocular 3D motion computations could be performed by

tracking an object's disparity-based *position-in-depth* over time; referred to as the changing disparity (CD) cue. Compelling evidence for such a mechanism was provided by Julesz & Payne (1968) in demonstrating that 3D motion percepts could be generated in the absence of coherent retinal motion signals. This can be accomplished by randomly relocating the elements of a moving dot stereogram on every display frame while maintaining steadily changing disparities over time; producing a percept similar to a detuned television screen moving through depth (detailed descriptions & demonstrations provided in Methods sections of Chapters 2 & 3). The observation that 3D motion percepts can be produced from purely cyclopean signals (i.e. only available after signals have been combined across the two eyes; Norcia & Tyler, 1984) has been a driving force in the belief that the CD cue is the primary cue to 3D motion (Cumming & Parker, 1994; Patterson, 2009; Regan & Gray, 2009).

However, as an object moves through depth, it not only creates changing binocular disparities over time, but also produces slightly different monocular velocities in the two eyes. In the extreme case—an object moving directly towards or away from the observer—equal and opposite directions of horizontal motion are projected on the two eyes. By computing interocular velocity differences (IOVD) in corresponding regions of the two retinae, the visual system could compute 3D directions of motion using a purely velocity-based mechanism. Although interest in the velocity-based cue has peaked in recent years, the potential for an IOVD computation was acknowledged in many of the earliest endeavors into binocular 3D motion processing (e.g. Beverley & Regan, 1973; Zeki, 1974). One reason for the delayed interest/exploration of the IOVD mechanism is that true isolation from concurrent changing disparity information is not as straightforward as was in the case of the CD cue. However, a number of experimental stimuli & methods have been developed to effectively remove (though not statistically eliminate) coherent changing disparity information while maintaining interocular velocity differences, including: binocularly *un*correlated moving dot stereograms (Shioiri, Saisho, & Yaguchi, 2000), binocular contrast anti-correlation (Rokers, Cormack, & Huk, 2008a), and induction of IOVDs by monocular motion adaptation (Brooks, 2002a; Fernandez &

Farell, 2006; Shioiri, Kakehi, Tashiro, & Yaguchi, 2009). Although there are pros and cons to each of these approaches, our preferred method of effective IOVD isolation is by anti-correlating dot contrast between the two eyes. By presenting binocularly paired dots with opposite contrast polarity between the two eyes (i.e. pairing a bright dot in one eye with a dark dot in the other), coherent disparity-based position information is severely degraded, while maintaining coherent monocular velocity signals necessary for IOVD computations (J. M. Harris & Rushton, 2003; Rokers, Cormack, & Huk, 2008a).

EXPERIMENTAL AIMS

The first step in exploring binocular 3D motion processing is to examine conditions in which the velocity- or disparity-based cue might be preferentially effective in signaling the direction of motion through depth. By considering the fundamental components of each binocular cue (static disparities or monocular velocities), we identified a range of conditions under which one might expect performance of the two cues to differ. Because the CD cue is based on comparisons of static disparity signals—which are known to be degraded at far eccentricities and high temporal frequencies (Blakemore, 1970; Cumming, 1995; Norcia & Tyler, 1984; Westheimer & Truong, 1988)—we hypothesized that the CD cue might be most effective for slowly moving stimuli presented near fixation. On the other hand, the IOVD cue relies on monocular velocity signals, which are robust across eccentricity and speed (McKee & Nakayama, 1984; Wright, 1987), and could be expected to be more heavily relied upon for higher speed stimuli presented outside the fovea. To test these hypotheses, we conducted extensive psychophysical measurements on the speed and eccentricity tuning of putative disparity- and velocity-based 3D motion mechanisms using a common metric of sensitivity; which allowed for direct comparison of sensitivities between motion conditions. Comparisons of the spatio-temporal tuning surfaces from isolated cue stimuli to that of a 'Full' cue stimulus, which contained both CD & IOVD cues (as in natural viewing conditions), provided an estimate of relative binocular cue contributions across a range of realistic viewing conditions.

In addition to what binocular 3D motion signals are computed, it is also important to understand how 3D motion directions are represented in the brain. Regardless of whether 3D motion is processed via disparity- or velocity-based mechanisms, binocular motion cues must be computed from the inherently two-dimensional patterns of light projected on the retinae. The additional computational overhead and the large ecological significance of 3D motion suggest that it may be selectively processed by the visual system. Motion adaptation paradigms are particularly useful in dissociating such intricately connected mechanisms (Alais, Verstraten, & Burr, 2005; van de Grind, Lankheet, & Tao, 2003; Yang & Lisberger, 2009). After viewing a unidirectional moving stimulus for a sustained period, subsequently viewed stationary stimuli appear to move in the direction opposite adaptation. This motion aftereffect has been taken as evidence for neurons tuned to the direction of frontoparallel (or retinal) motion (Grunewald & Lankheet, 1996; Hiris & Blake, 1992; Kohn & Movshon, 2003). Although there have been a number of motion aftereffect studies that employ 3D stimuli in some fashion (Brooks, 2002a; Fernandez & Farell, 2006; Shioiri et al., 2009), there is surprisingly little evidence to distinguish neural tuning for three-dimensional directions of motion (e.g. toward or away from the observer) from plausible combinations of 2D (monocular) directions of motion. The presence of a distinct 3D motion aftereffect would be convincing evidence that 3D motion directions are explicitly represented by the visual system. On the other hand, the absence of such an effect would suggest that 3D directions of motion may simply be readout from populations of neurons selective for the velocity and/or disparity components of 3D motion; as in readout mechanisms employed in fine vs. coarse perceptual decision making tasks (Jazayeri & Movshon, 2007).

In the third chapter, we test these two hypotheses using psychophysical motion adaptation paradigms to determine whether populations of neurons selective for 3D directions of motion exist. However, interpretation of a 3D motion aftereffect is not as straight forward as the 2D motion case. Because 3D motion generates distinct monocular velocities in the two eyes, a true 3D MAE must also be distinguished from simple combination of opposite monocular motion aftereffects. We address this by comparing

the magnitude of 3D MAE with aftereffects induced from monocular and frontoparallel motion stimuli.

If existence of a 3D motion aftereffect cannot be explained by inherited monocular motion adaptation, it would suggest that the 3D MAE arrises due to the adaptation in a mechanism specific to binocular motion signals. We therefore examined how binocular cues contribute to 3D motion adaptation by measuring adaptation to CD and IOVD isolating stimuli moving directly toward or away from the observer. Designed to closely align with previous spatio-temporal tuning experiments, we used the same cue isolating techniques as previously described. Crucially, this allowed us to measure 3D motion aftereffects at a single stimulus speed and eccentricity known to produce equal sensitivity in the two isolated cue conditions (i.e. the intersection of CD and IOVD spatio-temporal tuning surfaces). Providing matched drive to disparity and velocity based cues allowed us to hypothesize that resulting differences in the magnitude of isolated-cue motion aftereffects reflect an imbalance in the reliance on one binocular mechanism, rather than imbalance in adaptive drive.

To better understand how CD and IOVD computations fit into known motion processing pathways, we examine the physiological basis for distinct representation of 3D directions of motion using functional magnetic resonance imaging (fMRI) techniques. Previous work has implicated extrastriate areas in and around human MT+ (;) (Rokers, Cormack, & Huk, 2009; Likova & Tyler, 2007; respectively) as being involved in 3D motion processing, yet it is unclear whether this is the result of a unified '3D motion' mechanism or distinct disparity- and velocity-based mechanisms. Using event-related fMRI adaptation protocols (Huk, Ress, & Heeger, 2001; Larsson, Landy, & Heeger, 2006) we compare the the BOLD response following adaptation to isolated 3D motion cues to examine how the visual system might combine multiple sources of 3D motion information.

Chapter 2: Speed & eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion[1]

INTRODUCTION

Most research on motion perception has focused on two-dimensional (2D) frontoparallel motion (which is, of course, the easiest to generate on computer-driven displays). The neural computations that support the perception of motion through a more realistic 3D environment are considerably less well-understood. Here, we explore the relative contributions of two fundamental binocular cues to 3D motion (sometimes referred to as motion-in-depth or z-axis stereomotion): one based on changes in binocular disparity over time, and one based on interocular velocity differences[2]. We investigated the conditions under which the visual system might preferentially rely upon one cue over the other to shed light upon how 3D motion is processed by the brain.

The first potential cue, changing disparity over time (CD), is classically assumed to be the pure, fundamental signal for binocular 3D motion perception (Cumming & Parker, 1994; Gray & Regan, 1996; Regan & Gray, 2009). The CD cue can be computed simply by taking the time derivative of horizontal binocular disparity (Figure 2.1, upper). Although this cue is sufficient for the generation of 3D motion percepts, it is important to note that it is based on disparity signals, which are greatly degraded at far eccentricities and high temporal frequencies (e.g. Blakemore, 1970; Julesz, 1960; Norcia & Tyler, 1984; Westheimer & Truong, 1988). We therefore hypothesized that the CD cue might preferentially support 3D motion percepts for slowly-moving stimuli near fixation.

---

[1] Published work:
Czuba, T. B., Rokers, B., Huk, A. C., & Cormack, L. K. (2010). Speed and eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion. *Journal of Neurophysiology*, 104(5), 2886–2899. doi:10.1152/jn.00585.2009

[2] A percept of motion through depth has been shown to arise from moving objects visible to one eye and camouflaged to the other {Brooks:2006ff, Brooks:2007dm}. This is a dynamic analog of the stereopsis without binocular correlation described by Kaye {*Kaye:1978ti}. Although interesting, this is beyond the scope of the current paper.

The second potential cue, the interocular velocity difference (IOVD), has been proposed to also contribute to 3D motion perception (J. M. Harris, Nefs, & Grafton, 2008; but see Regan & Gray, 2009). The IOVD cue can be computed by comparing monocular velocity signals of the stimulus projections on each of the two retinae (Figure 2.1, lower). This computation exploits the fact that motion through depth projects different (and often opposite) motion upon the left and right eye's retinae. Although the IOVD cue has proven challenging to study in isolation, one might expect it to be perceptually robust across eccentricity and speed, given that it is based on monocular motion signals (e.g. McKee & Nakayama, 1984; Wright, 1987). We therefore hypothesized that the visual system may rely more heavily on the velocity-based IOVD cue to 3D motion for faster motions outside the fovea— the very conditions under which the disparity-based CD cue might be expected to be relatively ineffective.

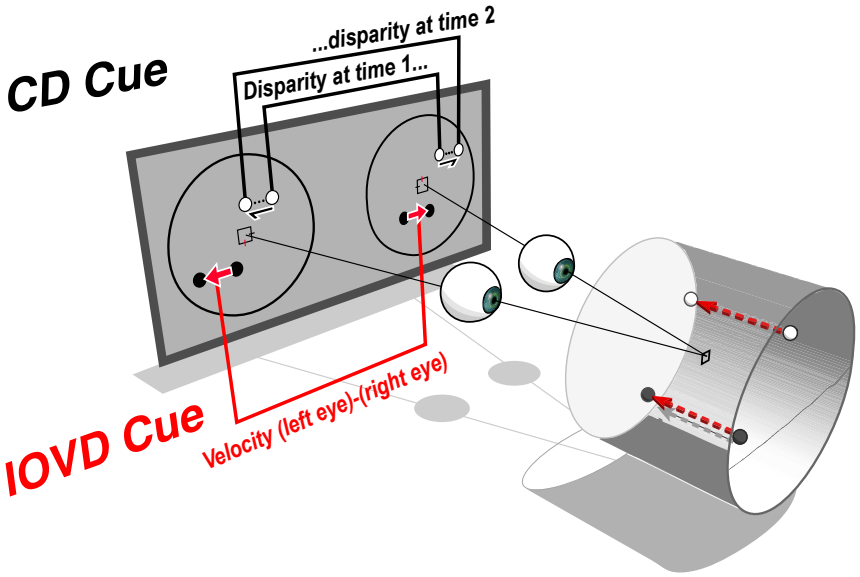

Figure 2.1 Schematic of binocular 3D motion cues

The schematic depicts two known binocular cues to 3D motion. The visual system could infer the direction of 3D motion (white and black spheres moving towards the eyes, in the cylinder at lower right) based on dynamic information from the corresponding 2D retinal projections (white and black dots, flat panel at upper left). The changing disparity (CD) cue is derived from tracking changes in binocular disparity over time (black brackets and ellipsis). The interocular velocity difference (IOVD) cue is derived by comparing monocular velocity signals from corresponding regions in the two retinae (red brackets and arrows).

In order to compare the relative contributions of the CD and IOVD cues, we assessed sensitivity with a common task and metric across all conditions. To do this, we asked observers to discriminate the direction of motion of a frontoparallel plane of dots moving towards or away through depth. We varied stimulus strength in a manner akin to prior manipulations of 2D motion coherence (i.e. the proportion of coherently-moving dots on the signal plane relative to noise dots moving incoherently through depth; Burr & Santoro, 2001; Lankheet, van Doorn, Bouman, & van de Grind, 2000; Newsome & Paré, 1988; Watamaniuk, Sekuler, & Williams, 1989). FULL binocular cue displays contained signal plane dots moving towards or away from the observer in depth (and hence contained both the CD and IOVD cues); CD displays contained 1-frame lifetime signal plane dots to remove any structured monocular velocity information (Julesz, 1971); IOVD displays contained binocularly-anticorrelated dots to greatly reduce the ability of the visual system to compute disparity (J. M. Harris & Rushton, 2003; Rokers, Cormack, & Huk, 2008a).

We measured direction discrimination thresholds for FULL, CD, and IOVD displays across a wide range of speeds and eccentricities. At low speeds, eccentricity degraded sensitivity similarly for all three display types. At faster speeds, however, the relative contributions of the two cues became much more distinct. For the FULL and IOVD stimuli, sensitivity showed a distinct bandpass speed tuning, with a peak at relatively fast speeds (~ 2.0°/sec monocular velocity). In contrast, sensitivity for the CD stimuli was clearly lowpass; in fact, direction discrimination for these stimuli became impossible at moderately high speeds, regardless of eccentricity.

These results imply that the visual system can compute 3D motion primarily from the IOVD cue across a majority of the visual field for a broad range of speeds, and may rely more strongly on the CD cue for direction discrimination at very slow speeds at or near fixation. Although our findings may at first seem to run counter to previous demonstrations that the CD cue is sufficient to explain 3D motion sensitivity (Cumming & Parker, 1994; Gray & Regan, 1996), our demonstration of the possible primacy of the IOVD cue bolsters a growing, recent literature demonstrating dissociable contributions of

the IOVD cue (for a more detailed review, see Discussion, and also Harris, Nefs, and Grafton 2008; Regan and Gray 2009). Furthermore, this relative primacy of IOVDs may result from the fact that our stimulus parameters, discrimination task, and sensitivity metric are better thought of extending the methods used to study 2D / frontoparallel motion processing into the third spatial dimension— as opposed to being temporally-dynamic extensions of the methods used to probe disparity-based mechanisms. Finally, the methods introduced in this paper could easily be generalized for use in future neuroimaging and electrophysiological studies of interocular velocity differences and 3D motion perception.

METHODS & MATERIALS

Observers

Data were collected in three experienced psychophysical observers (three of the authors, males aged 26–44), all with good stereopsis, and normal or corrected-to-normal vision. Experiments were undertaken with the written consent of each observer, and all procedures were approved by the UT-Austin Institutional Review Board. A total of 156,240 trials were collected across the 3 observers.

General procedure

We measured the ability of observers to discriminate the direction of motion through depth (directly towards or away from the observer) for three different types of motion cue stimuli (FULL, IOVD, or CD; described more fully below), which contained different combinations of the two primary binocular cues to 3D motion. Performance for each motion cue type was measured in a fully crossed design manipulating stimulus speed and eccentricity. Across all conditions, observers viewed a frontoparallel plane of random dots moving towards or away from them through a 3D volume of noise dots. The signal plane started at a random depth within the volume and moved at a smooth, constant speed either towards or away from the observer, wrapping from front-to-back (or

back-to-front) to complete one full cycle through the depth volume. The random starting location ensured that the starting, ending, or average disparity (or the time of the wrap) of the plane could not be used to do the task.

On each trial, observers viewed the display and reported the perceived direction of motion through depth (towards or away) with a left or right mouse click. The response triggered the next stimulus presentation with a minimum delay of 200 ms between trials. Observers were instructed to report their percept of the smoothest motion though depth throughout the experiment, disregarding the perceived jump through depth that occurred when the signal plane wrapped. In all conditions, signal dots were relocated in the x-y plane upon z-axis wrapping to minimize apparent motion during this brief change in signal plane depth. No feedback was provided; observers could thus concentrate on the smooth motion and not train themselves (consciously or not) on other potential cues, such as the jump due to wrapping, or a utrocular identification combined with a monocular direction judgment.

For each combination of cue type, speed, and eccentricity, we measured the proportion of correct responses as a function of motion coherence, defined as the relative percentage of signal to noise dots. Psychophysical thresholds (84% correct) were then estimated for each condition from a fitted logistic function, and sensitivity was expressed as inverse threshold coherence (coh-1). More details about the stimuli and experimental design are provided below.

General stimuli

Observers stereoscopically viewed moving random dot displays in which 80 dark (0.4 cd/m2) or light (129.7 cd/m2) binocularly paired dots were presented on a mid-gray (56.0 cd/m2) background (Figure 2.2, panel A). In each monocular half-image, half the dots were dark and half the dots were light. Individual dots subtended a visual angle of 15 arcmin (0.25°) and were anti-aliased to achieve subpixel position accuracy. Dot density and luminance values were not selected on a single fundamental principle, but rather a balance of several factors. First and foremost, we wanted our stimuli to be comparable to

many 2D frontoparallel motion studies that used motion coherence as a manipulation (i.e. number of dots on each display frame [80], stimulus area [126 deg2], and variable dot lifetimes, resulting in dot densities of 0.7–2.5 dots per degree and 3.7–11.5% dot coverage). We emphasize that many of our stimulus parameters may differ from prior studies of static disparity processing and stereomotion (e.g., lower density). Second, we chose parameters that were within hardware limitations (luminances that fit within the maximum linearized contrast range; dot numbers that did not overstep the available computational power to relocate each signal/noise dot on every display frame). Third, we chose parameters that allowed motion coherence manipulations to drive performance from chance to near perfect under all stimulus conditions.

Observers fixated a small central square (0.5°) with horizontal (black) and vertical (red) nonius lines. A single dot (bright, 0.25° diameter, 0 arcmin disparity) was placed in the fixation square to provide subjects with an object of fixation, and prevent fixation drift towards endpoints of the fixation square or nonius lines. To further aid in proper binocular alignment, four stationary dots (dark, 0.5° diameter, 10.6° eccentric, 0 arcmin disparity) were located beyond the stimulus on horizontal and vertical axes of each monocular half-image. We used a sparse set of reference dots to limit extraneous relative disparity cues at the outer edges of the display (Andrews, Glennerster, & Parker, 2001) while still providing eccentric visual anchor points.

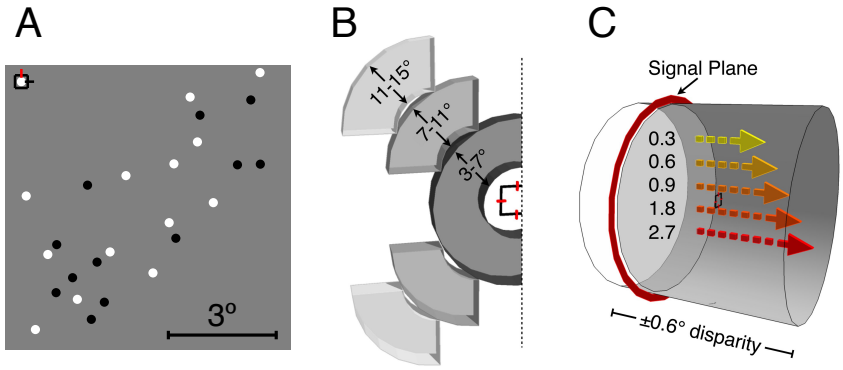

Figure 2.2 Experimental stimuli & manipulations

(A) Scale version of the stimulus as presented to the right eye. For clarity, this shows only the lower right quadrant of the monocular stimulus. The fixation point and nonius lines are located in the upper lefthand corner, with bright and dark stimulus dots scattered across an annular region 3–7º eccentric. Under all conditions the grey background extended across the entire monitor. By design, a single frame from one monocular image was identical (statistically) across all three motion cue stimuli. (B) Depiction of stimulus eccentricities. At the smallest eccentricity, signal and noise dots were restricted to an annular stimulus volume 3–7° from fixation. The smallest eccentricity was divided into 4 equal area quadrants (only the left 2 are shown, but the full display was left-right symmetric) that were displaced further outward for the other two eccentricity conditions, yielding eccentricities of 3–7°, 7–11°, and 11–15°. (C) Oblique depiction of the cyclopean 3D percept and speeds. A plane of signal dots moved through a cloud of noise dots (signal plane outlined in red for clarity, individual dots that actually constituted the plane not shown), and observers performed a 3D direction of motion discrimination (towards vs. away). The signal plane moved at 1 of 5 different speeds (colored arrows, in °/sec·eye), corresponding to motions through depth ranging from about 8 cm/sec to about 72 cm/sec (given our viewing distance of 70 cm). Direction discrimination sensitivity was measured at each of the 5 speeds (C) and 3 eccentricities (B), for each of the 3 motion cue types (see Figure 2.3).

Manipulations of eccentricity and speed

To examine how 3D motion sensitivity varies across the visual field, stimuli were presented within three different eccentricity ranges: 3-7°, 7-11°, and 11-15° from fixation (Figure 2.2, panel B). The 'Near' eccentricity stimulus consisted of a continuous annular region spanning 3-7° from fixation. From this annulus, four 90° annular segments were then shifted outward in oblique directions (45°, 135°, 225°, 315°) to 'Middle' and 'Far' eccentricities of 7-11° and 11-15°, respectively. Thus the number and density of signal dots were held constant across the eccentricity conditions. Stimulus disparities were constrained to a volume spanning ±0.6° of disparity (i.e., along the z-axis) from the plane of fixation. At our 70 cm viewing distance, this corresponded to a total (front to back) simulated depth interval of 16 cm. This z-axis depth, and hence the overall stimulus volume, remained constant across all conditions.

To examine how 3D motion sensitivity varies with stimulus speed, stimuli were presented at 5 different speeds (where we define "speed" as the monocular angular speed in each eye): 0.3, 0.6, 0.9, 1.8, and 2.7°/sec (Figure 2.2, panel C). Because we describe speed in °/sec per eye, and the monocular velocities were always opposite in the two eyes, one can simply multiply the monocular speeds times the number of eyes (2 in our case) to calculate the equivalent disparity change in °/sec.

Across all speeds, the total stimulus excursion through depth in a trial was always one full cycle through the stimulus volume (with a single "wrap" occurring on all trials, except those few in which the signal dots happened to begin at the very front or back of the volume). Stimulus presentations containing exactly one full cycle with a single depth wrap were chosen so that neither average depth (or disparity) over a trial, nor instantaneous depth (or disparity) at any point in the trial (e.g., starting or ending), could be used to perform the task. Given this one-cycle constraint, the resulting stimulus durations ranged from 2 s at the slowest speed to just over 0.2 s at the fastest speed (corresponding to 120 and 13 video frames, respectively). The decision to fix the total

depth traveled (and not the overall duration of motion) was supported by three factors. First, our main inferences are based on comparisons of sensitivity across FULL, CD, and IOVD conditions, which effectively balances duration across the comparisons of interest. Second, given that we observed peak sensitivity in the main experiment at rather fast speeds (and hence, at short durations), we are confident that shorter durations per se did not strongly impair performance. Third, exploratory manipulation of duration at the medium and high speeds revealed only a very small effect of stimulus duration which, in any event, was balanced across motion cue conditions in the main experiment.

Motion cue conditions: FULL, IOVD, and CD

Three motion cue stimuli were employed: FULL, IOVD, and CD. All three stimulus types contained a single plane of signal dots moving towards or away from the observer through depth, in the presence of noise dots (described further below). Figure 2.3 schematizes the three motion cue stimulus types.

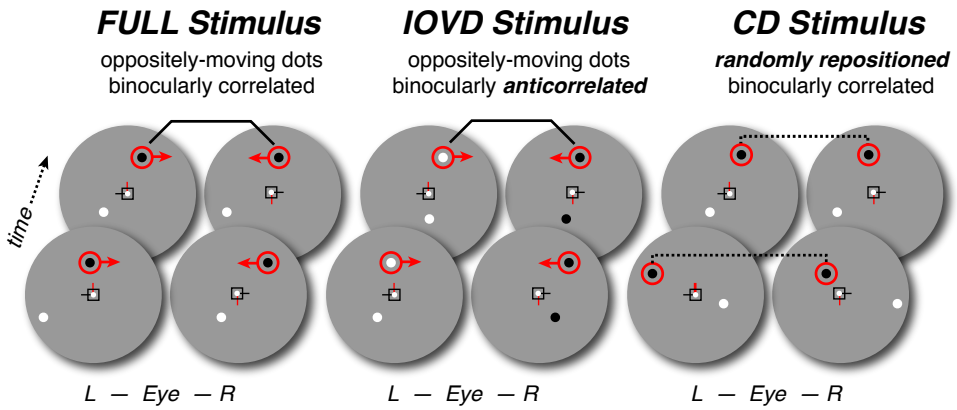

Figure 2.3 Isolation of binocular cues

Depiction of representative signal dots from two example frames from each of the three stimulus conditions (from left to right: FULL, IOVD, CD). The half images in each set can be free-fused (L, left eye view; R, right eye view). In the FULL stimulus, corresponding dots moved in opposite directions in the two eyes. Such a stimulus contains both changing disparities (the CD cue) and inter-ocular velocity differences (the IOVD cue). Red circles and arrows were not present in the actual stimulus, of course, but depict the respective motions of left and right eye's views of a single signal dot. Also note that actual dot densities were much higher (see Figure 2.2A and Methods): many similar signal dots specified a plane moving towards or away through depth. In the IOVD stimulus, corresponding dots also moved in opposite directions (just as in the FULL stimulus), but the dots were binocularly anticorrelated: a black dot in one eye was paired with a white dot in the other. This greatly reduced the contribution of the disparity-based CD cue, but the IOVD signal was preserved. In the CD stimulus, signal dots were randomly repositioned every frame, but their disparities (dotted brackets) still specified a signal plane moving towards or away through depth. The 1-frame lifetimes of the dots removed any coherent monocular motions and hence eliminated the IOVD cue, while preserving the CD signal.

The FULL stimulus consisted of a moving random dot stereogram in which binocularly-paired signal dots moved in opposite directions in the two eyes. The signal dots thus contained both the IOVD and CD cues to 3D motion. Signal dots moved at constant monocular speeds ranging from 0.3 to 2.7 degrees per second, corresponding to 3D motion speeds of about 8 cm/sec to about 74 cm/sec at our 70 cm viewing distance. Perceptually, the FULL stimulus resembled a fixed set of dots, not unlike a group of flying insects in a fronto-parallel plane, moving directly towards or away from the observer in synchrony. All dot pairs (signal and noise) were binocularly correlated (i.e., were of the same contrast polarity across the eyes); noise dots are described in detail later, and followed identical motion patterns across all three conditions.

The IOVD stimulus was identical to the FULL stimulus, except that the all of the dot pairs were binocularly anticorrelated: each dark dot in one eye was paired with a corresponding bright dot in the other eye. Anticorrelation has been shown to disrupt static disparity mechanisms (Cogan, Kontsevich, Lomakin, Halpern, & Blake, 1995; Cumming, Shapiro, & Parker, 1998; Neri, Parker, & Blakemore, 1999), while maintaining monocular velocity information (J. M. Harris & Rushton, 2003), and hence, retaining IOVDs in the presence of greatly-degraded disparity-based signals. In addition, our group has previously used sparse, anticorrelated dot displays to isolate the contribution of the IOVD cue (Rokers et al., 2009; Rokers, Cormack, & Huk, 2008a). Perceptually, the IOVD stimulus is phenomenologically rather interesting (although this is not the direct subject of the current paper). Because static disparity information is greatly compromised, one generally has the sensation of "something" moving towards or away from the eyes, with neither a firm sense of a distinct plane of dots present in space, nor a sense of the position-in-depth of these moving elements.

The CD stimulus was identical to the FULL stimulus, except that the signal dots were randomly replotted in new x-y positions on the signal plane upon each screen refresh (i.e. at 60 Hz). This replotting removed coherent monocular velocity information, while preserving steadily-changing disparity information (Braddick, 1974; Cumming & Parker, 1994; Julesz, 1971). For the CD stimulus, the rate of disparity change matched

that of the FULL and IOVD stimuli. Perceptually, the CD stimulus resembles a plane of TV snow moving through depth towards or away from the observer.

Manipulation of 3D motion coherence

3D motion coherence, defined as the ratio of signal dots to noise dots, was randomly varied on a trial-by-trial basis according to the method of constant stimuli. We determined direction discrimination thresholds in units of motion coherence from the resulting psychometric functions.

At the beginning of each trial, the number of signal dots (as determined by the coherence level pseudorandomly drawn for that trial) was selected (out of the 80 total dots). The remainder were designated as noise dots. The signal dots were randomly positioned on a single fronto-parallel plane moving towards or away from the observer. This signal plane began at a random position in depth, and moved throughout the entire depth range (wrapping when necessary), ending in the same position in depth as it began. Upon wrapping, each signal dot was assigned a new random x-y position in addition to moving to the opposite end of the volume on the z-axis. For a given speed, this implied that the signal dots followed a uniform distribution of lifetimes between 1 frame and the number of frames in a trial at that speed (i.e. 120 frames at the slowest speed, and 13 frames at the fastest speed). For any single trial, there were thus two signal dot lifetimes, one pre-wrap and one post-wrap, which summed to the total number of frames in that trial.

We designed our noise dots to satisfy multiple, somewhat competing demands: 1) remaining constant (statistically) across all conditions (cue type by speed by eccentricity); 2) being capable of effectively masking the motion through depth of the signal plane across all conditions (and thus allowing us to measure psychometric functions spanning the full range of possible performance across all conditions); and 3) effectively matching the spatiotemporal properties of the signal dots per se across all conditions— that is, we did not want the signal dots themselves to "pop out" in any

condition due to either flashing on for one frame (CD condition), or persisting for multiple frames that varied with speed (IOVD and FULL conditions).

To satisfy these constraints, we used variable-lifetime noise dots that approximated random walks though depth. Their instantaneous (frame-to-frame) velocity was variable along the z-axis, but was constrained to be less than or equal to the signal dot velocity through depth (their x-y positions were fixed throughout each lifetime). Each noise dot was assigned a random lifetime ranging from 1 to 12 frames (16.7 to 200 ms) from an inverse squared distribution. Specifically, the probability of a noise dot having a given lifetime, L (in frames) was proportional to 1/L2, with L = 1, 2, ..., 12. At the expiration of a noise dot's lifetime, it was randomly repositioned within the stimulus volume and assigned a new lifetime from this distribution (Our noise dots can therefore be considered a hybrid of the random-position and random-walk same-selection dot noise described in Scase, Braddick, & Raymond, 1996). At any given time, then, the noise consisted of a mixture of transient and persistent dots, like the CD signal dots on the one hand and the IOVD and FULL signal dots on the other, but with the distribution favoring the presence of short lifetime dots.

The distribution of noise dots included a higher proportion of shorter lifetimes, because we reasoned that transient (flashing) elements are better at masking persistent elements than vice versa. Noise composed of mostly transient elements would be expected to mask both transient and persistent signals (which is important, given that our CD condition contained transient signal due to the single-frame lifetimes of the signal dots, and our FULL and IOVD conditions contained more persistent signal due to the longer signal dot lifetimes). This argument can also be appreciated in the Fourier domain: transient noise elements will have broadband power in the temporal frequency domain, and thus would cover the spectral range of signal across all conditions; persistent noise elements would be better localized in the temporal frequency domain, and thus would not so broadly span the spectral range of signals of interest. Finally, pilot observations confirmed that noise from an inverse squared distribution yielded good subjective degradation of the signal plane's direction of motion at high noise levels, provided good

masking of the dots themselves across conditions (i.e., none of the signal dots popped out in any condition), and drove performance from an upper asymptote to chance levels for all motion cue types. We emphasize that these decisions allowed us to use noise dots that had the same distribution of motions and were subjected to the same manipulation of coherence (and hence, the same sensitivity metric) across all conditions; a crucial component that enabled direct comparison of sensitivities across cue conditions.

Experimental design

We measured observers' ability to discriminate the direction of motion through depth (towards or away) across a range of 3D motion coherence levels (0, 3, 6, 12, 24 and 50% coherence) using the method of constant stimuli. We employed a fully-crossed design containing all combinations of motion cue type (FULL, CD, IOVD), eccentricity (3-7°, 7-11°, and 11-15°), and speed (0.3, 0.6, 0.9, 1.8, and 2.7°/sec·eye). Within each run, we measured percent correct as a function of motion coherence for a single combination of these factors (resulting in a single estimate of the psychometric function).

Motion coherence was pseudo-randomized across trials within a run. Each run consisted of 40 trials per coherence level, resulting in 240 trials total. The order of runs was randomized. Each observer completed five runs of the 0.6, 0.9, and 2.7°/sec·eye speeds for each motion cue/eccentricity combination, and three runs for each of the 0.3, and 1.8°/sec·eye speeds. This resulted in either 720 or 1200 trials per observer per condition, and just under 45.4 kilotrials per observer for the main experiment. Two control experiments addressing position-in-depth and 2D motion discrimination (see Discussion) contributed an additional ~20.2 kilotrials across the same three observers.

Apparatus and display

To investigate 3D motion perception at large eccentricity with high temporal accuracy, stimuli were presented on a calibrated 42" LCD Display (Sharp LC-42D64U; 60Hz progressive scan, 1920 x 1080 pixel resolution) viewed through a mirror

stereoscope with a concomitantly large field of view. The monitor was driven by Mac Pro computer and an NVIDIA GeForce 8800 GT video card.

Luminance calibrations were done at 10 locations across the display using an OptiCal photometer (CRS Ltd.). We verified that gamma correction tables for each location were the same across the entire luminance range, allowing all stimuli to be accurately presented using a single linearizing gamma correction table. All 10 curves were nearly identical within a scale factor, demonstrating a high degree of spatial luminance homogeneity, nearly perfect contrast homogeneity, and the ability to implement good luminance linearization with this LCD display.

We achieved spatial luminance homogeneity by making internal display adjustments. Specifically, the duty-cycle of the LCD backlight was maximized by setting the "backlight" adjustment to the maximum level, while setting the "brightness" adjustment to the minimum level to maintain a comfortable luminance range. This provided the most homogenous display luminance, leaving at most a 10% residual luminance variation that was almost entirely constrained to the extreme edges of the display (where stimuli were not presented).

Display timing was verified using a fast photocell (Model 10AP, UDT Sensors Inc., Hawthorne, CA) and an oscilloscope. We used a splitter-cable so that we could measure the VBL signal directly while simultaneously measuring the instantaneous luminance on the monitor. Pixel updates were constant at 60 Hz and at a fixed phase relative to the VBL signal generated by the video card. The white-to-black transition was marginally faster than the black-to-white, with the later showing slight exponential characteristics. Nonetheless, the display easily followed repeating black-white and black-gray-white-gray cycles on a frame-by-frame basis at 60 Hz. Although the display was slow by modern CRT standards, it provided reliable timing of display updates, as well as luminance output that was easily linearized. Furthermore, all in-monitor enhancement modes (e.g. motion enhancement, dynamic contrast adjustment, etc.) were disabled, as they could yield undesirable display artifacts. In short, our measurements suggested that our particular LCD was appropriate for use in our experiments; it remains to be seen

whether similar results can be attained by similar adjustment and calibration using other LCDs.

Monocular half-images were presented separately on the left and right halves of the display, with a septum and various baffles positioned to assure that each half-image was only visible to the corresponding eye. Viewed through the 70 cm optical path length of the stereoscope, each monocular half-image subtended 30° of visual angle. This display arrangement was selected over traditional dual-display stereo or shutter goggle apparatus because it provided both perfect temporal synchronization between the two eyes and complete isolation of the monocular half-images. All stimuli were generated using the Psychophysics Toolbox (Brainard, 1997) and MATLAB (2007a, The Mathworks Inc., Natick, MA).

Data analysis

For each condition (motion cue type x eccentricity x speed), we combined data across multiple runs for each subject, and fit a logistic psychometric function using the psignifit toolbox version 2.5.6 for Matlab (http://bootstrap-software.org/psignifit/). Threshold was defined as the 3D motion coherence yielding 84% accuracy. We bootstrapped confidence intervals (equivalent to ± 1 SEM) about these thresholds by resampling (with replacement) the binomial responses from each subject to create 500 repetitions of the experiment, fitting a psychometric function to each resampled experiment, and identifying the central 68% of the values. In instances where observers were unable to discriminate 3D motion direction we assigned a threshold coherence of 100% (pinning the thresholds at the maximum physically-realizable level was preferable to simply discarding those data, but our conclusions do not change if these are instead omitted). Across a total of 135 sensitivity estimates (3 observers, 5 speeds, 3

eccentricities, 3 cues) this occurred only 7 times, and was isolated to high speed CD stimulus conditions[3].

We applied a similar resampling approach when fitting the eccentricity and speed tuning curves (Figures 2.4 and 2.5). We plotted the median fit parameters (after checking that the median values were very similar to the means), because the medians had the advantage of being robust to the occasional extreme values that can arise in a small number of fits across the very large number of resampled datasets.

RESULTS

Recall that, in all conditions, observers simply judged whether a plane of signal dots was moving towards or away from them. On each trial, stimuli were presented at one of six different motion coherence levels (0, 3, 6, 12, 24 and 50% coherence). Sensitivity (inverse 3D motion coherence threshold) was then estimated for each combination of eccentricity (Near, Middle, and Far), speed (0.3, 0.6, 0.9, 1.8, and 2.7°/sec·eye), and motion cue type (FULL, CD, and IOVD).

In each of the following sections, we first establish a baseline for 3D motion discrimination by describing the performance in the FULL cue condition, which contained both disparity- and velocity-based 3D motion cues (CD plus IOVD). We then compare the results to each isolated cue condition (CD, IOVD). In order to explore how sensitivity varied across the entire eccentricity-speed space, we address the results from three perspectives in the following three sections: the effects of eccentricity at different speeds, the effects of speed at different eccentricities, and finally, the full spatiotemporal (speed x eccentricity) sensitivity surface. (Unless otherwise noted, data points in the following figures represent mean sensitivity across all three observers).

---

[3] Specifically, observer ACH was unable to discriminate CD stimulus direction at three highest speeds in the farthest eccentricity conditions (11° eccent.; 0.9, 1.8, and 2.7°/sec·eye) and at the highest speed in the middle eccentricity condition (7° eccent.; 2.7°/sec·eye). Observer LKC was unable to discriminate CD stimulus direction at any eccentricity at the highest speed (3, 7, 11° eccent.; 2.7°/sec·eye).

Effects of eccentricity at different speeds

Here, we first describe the data as functions of eccentricity measured at different speeds. Figure 2.4 (left) shows the eccentricity effect on FULL stimulus sensitivity across the range of 3D motion speeds. Increases in stimulus eccentricity caused a decrease in direction discrimination sensitivity in a speed dependent manner. The strength of this effect can be determined from the slope of linear fits at each stimulus speed. Sensitivity to slower 3D motion (darker curves and symbols) was strongly diminished by increasing eccentricity (slope of –0.83 coh–1·deg–1 at slowest speed, 0.3°/sec·eye), while sensitivity for faster 3D motions (lighter gray curves and symbols) was not strongly affected by eccentricity (slope of 0.02 coh–1·deg–1 at fastest speed, 2.7°/sec·eye). The eccentricity effect was smaller at the highest speed primarily due to a large improvement in sensitivity at the Far eccentricity; sensitivity at the Near eccentricity did not change as much. In summary, eccentricity reduced direction discrimination sensitivity in the FULL condition, but did so strongly for slower speeds, and less so for faster speeds.

Sensitivity to the IOVD stimulus (Figure 2.4, middle) showed a similar pattern of eccentricity dependence as the FULL stimulus. Sensitivity to IOVD motion was strongly affected by eccentricity at slower speeds (slope of –0.60 coh–1·deg–1 at the slowest speed, 0.3°/sec·eye), and was less affected by eccentricity at higher speeds (slope of –0.16 coh–1·deg–1 at the fastest speed, 2.7°/sec·eye). As with the FULL stimulus, the change in eccentricity function across speeds was more due to changes in sensitivity at Far eccentricity than at Near eccentricity. So, just as for the FULL condition, eccentricity effects for the IOVD condition were larger at slower speeds, and smaller at higher speeds.

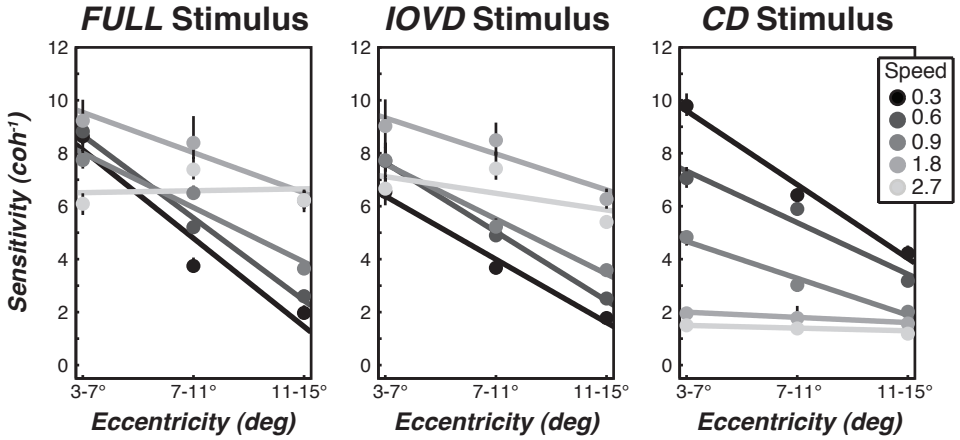

Figure 2.4 Effects of eccentricity at different speeds

Direction-discrimination sensitivity (y-axis) as a function of eccentricity (x-axis), with speed as the grouping parameter (lighter shades of gray corresponding to faster speeds). Error bars represent 68% bootstrapped confidence intervals (i.e. standard errors). For both the FULL (left) and IOVD (middle) conditions, sensitivity decreased with increasing eccentricity. Moreover, sensitivity generally increased with increasing speed for these two conditions, particularly at higher eccentricities. Only for the very fastest speed did sensitivity begin to decrease (this effect was most pronounced at the smallest eccentricity). The pattern in the CD condition (right) was strikingly different: sensitivity did decrease with increasing eccentricity and more so for slowest speeds, but the overall order of the curves was reversed, with low speeds yielding much higher sensitivities than high speeds.

Sensitivity to the CD stimulus (Figure 2.4, right) followed a strikingly different pattern than that seen from the FULL and IOVD stimuli. Larger eccentricity did yield poorer performance in general, and eccentricity effects were steepest at the slower speeds . At the higher speeds, however, performance was very poor regardless of eccentricity. At the highest speed, in fact, only one observer was able to reliably perform above chance. In other words, the lack of an eccentricity effect at high speeds is probably best thought of not as a uniform sensitivity across eccentricities per se, but rather as an overall lack of sensitivity of the CD system to stimuli moving rapidly in depth.

In summary, the effects of eccentricity on FULL and IOVD sensitivity were quite similar, and showed similar dependencies on speed. In contrast, CD sensitivity followed a very different pattern of interactions between eccentricity and speed. The nature of this interaction becomes more clear in the next section.

Effects of speed at different eccentricities

It is perhaps more illuminating to consider the same sensitivity data as speed-tuning curves measured at different eccentricities, as shown in Figure 2.5 (note the log speed axis). For the FULL stimulus (Figure 2.5, left) similar bandpass speed tuning functions were evident at all stimulus eccentricities; peaking near the higher speeds measured. At Near eccentricity, sensitivity fell off sharply for speeds faster than 1.8°/sec·eye. Because of the bandpass appearance, we fit the data with Gaussian functions. The peak of the fitted Gaussian for the Near data was at 1.09°/sec·eye, with a full width at half the maximum height (FWHM) of 3.92°/sec·eye. At Middle and Far eccentricities, peak sensitivity shifted towards even higher stimulus speeds (Middle,1.94°/sec·eye; Far, 2.27°/sec·eye). The stronger effects of eccentricity at low speeds had the effect of narrowing the bandpass speed tuning at Middle and Far eccentricities (FWHM, Middle, 3.23°/sec·eye; Far, 3.12°/sec·eye).

The pattern of IOVD sensitivity (Figure 2.5, middle) was again strikingly similar to the FULL stimulus. Sensitivity was bandpass for speed with a peak near the higher speeds. This bandpass tuning was evident at all eccentricities, and IOVD peak sensitivity

also followed a similar pattern as for the FULL stimulus. Peak speeds were 1.54, 2.06, 2.06°/sec·eye for Near, Middle, and Far eccentricities respectively, and likewise became more sharply speed-tuned with increasing eccentricity (FWHMs were 3.49, 3.23, and 2.91°/sec·eye, respectively). The similarity between the FULL and IOVD patterns of speed tuning is further supported by a point-by-point comparison of the two tuning functions; revealing that twelve of the fifteen points fall within 68% (± 1 SEM) confidence intervals of one another.

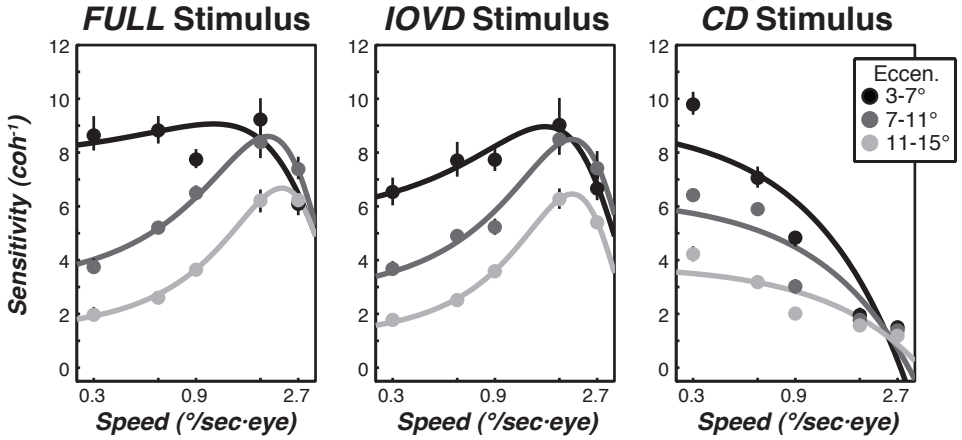

Figure 2.5 Effects of speed at different eccentricities

Direction-discrimination sensitivity (y-axis) as a function of speed (x-axis), with eccentricity as the grouping parameter (lighter shades of gray corresponding to larger eccentricities). The speed axis is logarithmic, error bars represent 68% bootstrapped confidence intervals. For both the FULL (left) and IOVD (middle) conditions, sensitivity shows distinct bandpass tuning, with peak sensitivity occurring just before the highest speeds tested. In contrast, sensitivity in the CD condition (right) exhibited clear lowpass tuning, with maximal sensitivity for the speed closest to stationary. For all conditions, sensitivity generally increased with decreasing eccentricity (dark curves above light curves). Although at faster speeds sensitivity for FULL and IOVD conditions was generally higher than for the CD condition, note also that at the lowest speeds, sensitivity for the CD condition actually exceeded that of the FULL cue condition.

The CD condition demonstrated an altogether different pattern of sensitivity from those seen in the FULL or IOVD conditions (Figure 2.5, right). All of the speed tuning curves were clearly lowpass, falling off precipitously with increasing speed (i.e., only one observer was able to discriminate the highest speed CD stimuli at accuracies above chance). In fact, the difference in the tuning curves was pronounced enough that we couldn't fit the CD data satisfactorily with Gaussians given reasonable parameter values, and we therefore fit them with straight lines. Fitted linear slopes were all strongly negative (in units of sensitivity, coh–1 per °/sec: Near, –3.28; Middle, –2.09; Far, –1.10). As with the eccentricity effects in the previous section, the slopes of the speed effects on the CD stimulus became less steep at far eccentricities simply because accuracy levels fell towards chance.

In summary, the analysis of speed tuning reveals that IOVD-based performance closely mirrored that of full-cue performance. FULL and IOVD speed tuning was bandpass, with a peak at relatively brisk 3D motion speeds. In contrast, CD-based performance exhibited dramatically different speed tuning that cannot account for most of the full-cue sensitivity. Instead, CD speed tuning was lowpass, and fell off steeply near the speeds at which the IOVD and FULL conditions revealed maximal sensitivity. Although the stimuli and tasks were different, these speed tuning results are qualitatively consistent with the temporal frequency tuning results of Shioiri et al. (2008).

Speed by eccentricity (SxE) sensitivity surface

The preceding sections show that IOVD sensitivity is very similar to FULL sensitivity, and that CD sensitivity follows a rather different pattern, regardless of whether the data are viewed as slices of constant speed or constant eccentricity. The overall shape of the data for the three conditions can be appreciated more thoroughly in spatiotemporal sensitivity surface-contours that span both eccentricity and speed (Figure 2.6). The top row depicts the sensitivity surface for each motion cue condition; stimulus eccentricity by speed (SxE) on the x- and y-axes respectively, and direction discrimination sensitivity on the z-axis (height). The surface sensitivities are also

projected down to contour maps on the z=0 plane. Bandpass speed sensitivity can be seen in both FULL and IOVD conditions, as can the weaker eccentricity effects closer to the best speed. In contrast, the CD condition shows a distinct pattern of roughly linear sensitivity falloff as speed and eccentricity increase.

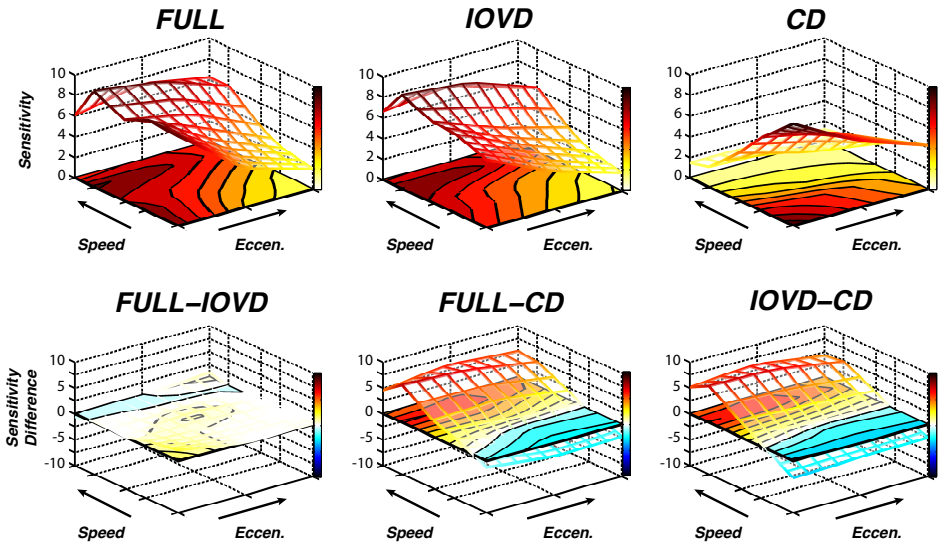

Figure 2.6 Spatiotemporal sensitivity surfaces

3D motion direction-discrimination sensitivity as a function of both speed and eccentricity. Top row, sensitivity as a function of both speed and eccentricity for the three cue conditions. Each sensitivity surface represents combined subject data from 45,360 trials. Height of mesh (z-axis) indicates sensitivity, as a function of speed and eccentricity (x and y axes). Colored floor is a contour plot depicting the same sensitivity information. The FULL and IOVD surfaces are quite similar, whereas the CD surface is distinctly different (so much so that, from this perspective, the view is of the bottom of the mesh surface). Bottom row, differential sensitivity surfaces highlighting the similarity, or lack thereof, between the three conditions (z=0 plane has been raised to allow for negative values). The FULL–IOVD surface is nearly flat, indicating substantial similarity. In contrast, the other two surfaces (FULL–CD, IOVD–CD) show a systematic pattern of differences. At the slowest speed tested, CD sensitivity was higher than IOVD sensitivity and higher than the FULL sensitivity as well (cool colored mesh and surface).

The bottom row of surface-contour plots (Figure 2.6) show the differential sensitivity surfaces generated by subtracting the spatiotemporal sensitivity surfaces of each motion cue. Positive values are shown in the same warm color map as the original sensitivity surfaces, while negative values are shown in cool colors extending below the contour map. The left two panels of the bottom row of surface-contour plots show differential sensitivity surfaces generated by subtracting each isolated cue surface from the FULL stimulus surface (i.e., FULL–IOVD, FULL–CD). The surface goes positive (above the contour map) where sensitivity is better in the FULL stimulus, and the surface goes negative (below the contour map) where sensitivity is better in an isolated cue condition (IOVD or CD). Bootstrapped p-values for the difference surfaces are shown in Table 2.1.

Table 2.1. P-values on 3D motion sensitivity difference surfaces.

Speed	Near	Middle	Far
		FULL–IOVD	
0.3°/sec·eye	0.004*	0.416	0.278
0.6°/sec·eye	0.090	0.200	0.322
0.9°/sec·eye	0.476	0.002*	0.416
1.8°/sec·eye	0.436	0.462	0.462
2.7°/sec·eye	0.162	0.460	0.034
		FULL–CD	
0.3°/sec·eye	0.066	0.000*	0.000*
0.6°/sec·eye	0.012*	0.020*	0.006*
0.9°/sec·eye	0.000*	0.000*	0.000*
1.8°/sec·eye	0.000*	0.000*	0.000*
2.7°/sec·eye	0.000*	0.000*	0.000*
		IOVD–CD	
0.3°/sec·eye	0.000*	0.000*	0.000*
0.6°/sec·eye	0.198	0.004*	0.002*
0.9°/sec·eye	0.000*	0.000*	0.000*
1.8°/sec·eye	0.000*	0.000*	0.000*
2.7°/sec·eye	0.000*	0.000*	0.000*

* $P < 0.05$

The FULL–IOVD surface shown in the lower left panel is nearly flat, indicating that the sensitivities are nearly identical at each combination of speed and eccentricity tested. In fact, only two of the fifteen points in the FULL–IOVD surface are significantly different from zero.

The FULL–CD surface shown in the lower middle panel, however, is distinctly not flat. For faster speeds, observers are much more sensitive to the FULL stimulus than the CD stimulus, indicating that the (lack of) CD sensitivity does not limit the observer's performance. For the slowest speeds, on the other hand, observers actually performed better in the CD condition than they did in the FULL condition, indicating that: (1) observers were unable to fully exploit the changing disparity information when interocular velocity differences were also present; and/or (2) there was richer changing-disparity information present in the CD stimulus than in the FULL stimulus (perhaps due to the faster temporal refresh rate of the signal dots in the CD condition). Regardless of which possibility is at work (both could be), this superiority of the CD condition for slow speeds demonstrates that the larger-scale dissimilarity between the FULL and CD conditions is not simply because the CD stimulus did not contain strong signal.

The lower right surface-contour plot shows the differential sensitivity surface generated by subtracting the CD sensitivity surface from the IOVD sensitivity surface (IOVD–CD). Obviously, this difference surface closely resembles the FULL–CD surface, and it can also be thought of as a visualization of the relative utility of the two (isolated) cues in our experimental conditions.

The ability of each isolated cue to predict the FULL sensitivity is shown in Figure 2.7. Individual observer sensitivities for each isolated cue condition (y-axis) are plotted as function of FULL stimulus sensitivity (x-axis). Each data point, in other words, shows a pair of thresholds for a particular combination of observer, speed, and eccentricity. The left scatter plot shows a high level of correlation between the IOVD and FULL stimulus sensitivities ($r2 = 0.75$). This strong correlation on an individual subject level suggests that FULL sensitivity can be accurately predicted simply by measuring an individual's corresponding IOVD sensitivity in isolation. However, this relationship does not hold for

the right scatter plot of CD vs FULL sensitivities (r2 = 0.05). Thus, knowing an individual's CD sensitivity does not provide much information to predict how well the observer will be able to discriminate the direction of realistic (full-cue) motions through depth. Interestingly, Watanabe et al. (2008) found a very similar pattern of results when comparing their novel clinical test for motion through depth with a standard static stereo test (Titmus).

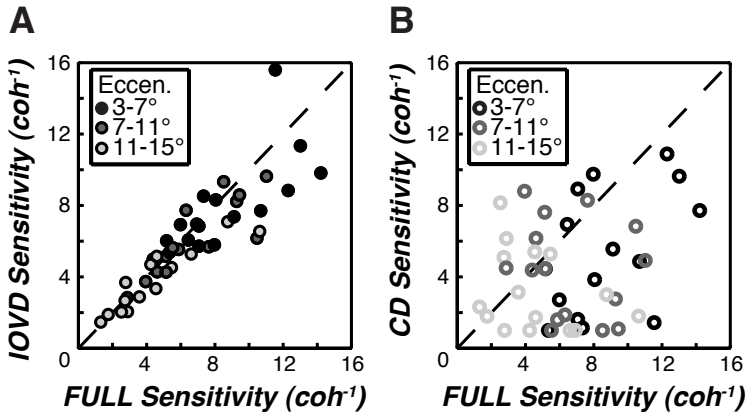

Figure 2.7 Correlation of isolated and combined cue sensitivity

Scatterplots show IOVD sensitivity (A) and CD sensitivity (B) plotted on the y-axis against corresponding FULL sensitivity (x-axis). Each data point corresponds to the sensitivity of an individual subject for a particular eccentricity/speed/motion cue condition. Lighter gray symbols represent farther eccentricities. The dashed line shows unity. IOVD sensitivity generally matched FULL sensitivity (r2 = 0.75), while CD sensitivity was relatively unrelated (r2 = 0.05), illustrating FULL sensitivity is better predicted by IOVD sensitivity than CD sensitivity.

DISCUSSION

Our experiments revealed distinctly different patterns of sensitivity to the changing disparity (CD) and interocular velocity difference (IOVD) cues for 3D motion direction discrimination. Sensitivity to the CD cue was highest at the shortest eccentricities and the slowest speeds. Increasing either speed or eccentricity had strong and independently deleterious effects on CD sensitivity. Sensitivity to the IOVD cue, on the other hand, was lowest at the nearest eccentricities and the slowest speeds. Increasing speed led to greater IOVD sensitivity and mitigated the effects of eccentricity. Overall, the pattern of IOVD sensitivity was nearly identical to the pattern of FULL sensitivity— sensitivity for stimuli containing both the CD and the IOVD cues— across the entire eccentricity-speed space. In contrast, the pattern of CD sensitivity across the eccentricity-speed space was markedly different than the FULL pattern. Although these patterns of relative sensitivity were not as straightforwardly dependent on speed and eccentricity as we initially hypothesized, they did reveal a surprisingly close correspondence between IOVD and FULL sensitivity across the majority of the wide spatiotemporal range we investigated. We therefore conclude that, at least outside the fovea, the human visual system can rely primarily on interocular velocity differences— not changing disparities— to discriminate the direction of 3D motion.

Distinguishing the contributions of the CD and IOVD cues

The differential patterns of sensitivity across the eccentricity-speed space provide clear evidence for a dissociation of the CD and IOVD cues. The disparity-based cue functions best at slow speeds and nearer eccentricities, while the velocity-based cue exhibits bandpass sensitivity for higher speeds, with muted eccentricity effects. These distinct patterns of sensitivity suggests that the CD cue may be useful for slow-moving (para-)foveal 3D motions, and that the IOVD cue may be more useful for faster and more peripheral 3D motions.

Our results also provide additional support to the notion that the IOVD cue can be experimentally isolated. We addressed the effectiveness of anticorrelation in isolating the

IOVD cue by conducting two experimental controls: (1) to determine whether useful position-in-depth information could be contributing to 3D motion sensitivities in the IOVD stimuli, and (2) to rule out the possibility that similarities in FULL and IOVD sensitivities could result from simple monocular motion sensitivities.

Unlike the CD cue, the IOVD cue cannot be perfectly isolated in principle. Although contrast anticorrelation is not a perfect form of isolation, we believe it is currently the most effective method of removing useful disparity information from a stimulus (and thus strongly biasing the stimuli in favor of IOVD mechanisms). There are two primary concerns that are often raised about these anticorrelated stimuli. The first is that, in an anticorrelated stimulus, there are many potential "false" matches of the same contrast polarity (e.g. a given white dot in one eye could conceivably be matched with any other white dot in the other eye's image, even though the experimenter had specified that the "corresponding" dot be of opposite contrast polarity). There several reasons why it is unlikely that these potential unintended matches influenced the data in a meaningful way. First, given the dot density of our stimuli, any unintended matches would have a large disparity, usually both horizontal and vertical, and these would vary from match to match at any given time (thus, a vertical vergence movement, for example, couldn't suddenly create a large number of plausible, predominantly horizontal disparities). As the effective signal of a binocular element falls off with the overall disparity, regardless of how "effective signal" is determined, the majority of these matches would not be a very effective stimulus (Blakemore, 1970; Cormack, Stevenson, & Schor, 1993; Prince, Cumming, & Parker, 2002; Stevenson, Cormack, & Schor, 1994; Stevenson, Cormack, Schor, & Tyler, 1992). Second, at threshold values of coherence, the vast majority of these potential matches for a given signal dot in one eye would be with a noise dot in the other eye. The match would thus result in an additional binocular noise dot (or at least a much noisier signal dot). If performance in the IOVD conditions were based on these spurious disparity signals, then it would be quite poor indeed. Rather, we find that the data from the IOVD conditions tracks that from the FULL conditions, despite the huge difference in the quality of the disparity signals. Even under the most general

assumptions about the presence of an unintended CD signal in our IOVD stimulus, our data are inconsistent with this explanation.

The second potential problem with anticorrelated stimuli is that each dot contains two vertical (on average) edge segments of opposite contrast polarity, so it is conceivable that, for example, the left edge of a white dot could be paired with the right edge of the corresponding black dot. However, these matches would be between regions of different overall (signed) contrast and local mean luminance, and it is known that unequal contrast between corresponding elements in the two eyes impairs stereopsis very dramatically—much more so than an overall contrast reduction (Cormack, Stevenson, & Schor, 1991).

However unlikely, these concerns are valid in principle, so we addressed them by conducting a control experiment where observers performed coarse position-in-depth judgments (2AFC discrimination of the signal dots as near or far relative to the plane of fixation), while viewing stimuli moving at 0.9°/sec·eye under each motion cue/ eccentricity combination. The position-in-depth judgments were performed on stimuli nearly identical to those of the main experiment except that one eye's image was flipped horizontally, so that the motion was in the same direction in the two eyes. This created moving (frontoparallel) stimuli with a fixed, random disparity offset in each trial while still maintaining the same monocular motions as in the main experiment.

When observers were asked to discriminate whether this plane of dots was near or far relative to the zero-disparity plane of fixation, performance with the anticorrelated (IOVD) stimulus was so poor that we could not measure psychometric functions and report thresholds. We therefore did the following analysis: For each eccentricity and observer, we noted the 84% correct threshold 3D motion coherence for the FULL stimuli. We then measured position-in-depth performance for each observer and eccentricity at this coherence for each stimulus type. So if, for example, there was rich disparity information extracted from the anticorrelated stimulus (due to an early rectifying nonlinearity, say), then performance in the disparity-based position-in-depth task should be close to 84% correct. If, on the other hand, the anticorrelated stimulus does indeed

greatly reduce the available disparity information, performance should be much poorer than 84% correct.

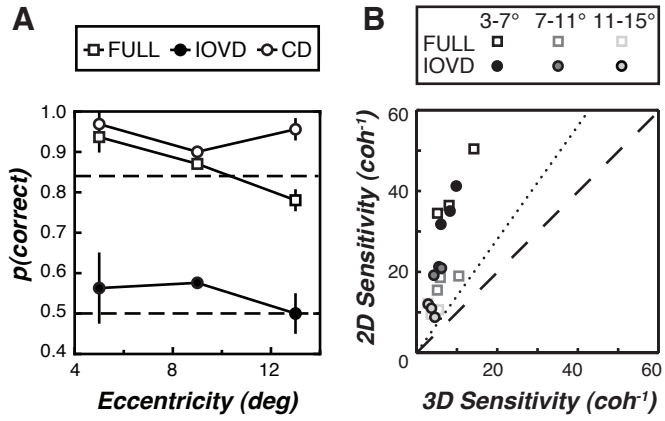

Figure 2.8 Confirmation of IOVD cue isolation

(A) Position-in-depth performance (proportion correct for a near vs. far 2AFC discrimination) for the CD (open circles), IOVD (closed circles), and FULL stimuli (open squares). Each point represents the position-in-depth performance (averaged over 3 observers) as measured at the motion coherence corresponding to each observer's 3D motion direction-discrimination threshold. CD position-in-depth performance was much better than IOVD at all eccentricities (and even better than that for the FULL stimulus in one case). Crucially, CD performance was always better than 84% correct (upper dashed line indicates the performance level for the 3D motion task at the coherences used), but IOVD performance was at or near chance (lower dashed line). (B) Individual observer's 2D motion direction-discrimination sensitivity for the FULL (open squares) and IOVD (closed circles) stimuli as a function of the corresponding 3D motion direction-discrimination sensitivity. All of the points fall above the dashed line (unity), indicating the much greater sensitivity to 2D frontoparallel motion than 3D motion. The dotted line (root-2 improvement 2D vs. 3D) suggests that the greater sensitivity to 2D motion cannot be explained by simple within-direction binocular summation.

Panel A of figure 2.8 is a plot of of the mean performance across subjects in the position-in-depth task for each cue type as a function of eccentricity, and tested at the corresponding 3D motion threshold coherence. The subjects all behaved very similarly (error bars are ± 1 S.E.M. across the 3 subjects, and are smaller than the symbols for 4 of the 9 points). Two trends are clear. First, (static) depth judgments for the CD stimuli at the 3D motion coherence threshold are better than 84% correct, confirming the rich disparity signals present in these stimuli. Second, and more importantly for our purposes, performance for all observers at all eccentricities was at or near chance for the IOVD stimuli, even though the same stimulus coherences supported 84% correct performance on the 3D motion task. This indicates that, what ever the potential disparity information is in the anticorrelated stimuli, observers were unable to use it to do basic depth discriminations.

The near-absolute failure to accurately judge position-in-depth on IOVD stimuli implies that any disparity signals arising from the anticorrelated elements were not perceptually accessible for the purposes of performing a simple near-vs.-far task. These observations support the notion that direction-discrimination of our IOVD stimulus in the main experiment was based on the interocular comparisons of velocities—with effectively no contribution of residual disparity signals that might have been used to compute an additional CD signal. This control experiment replicates and expands our previous dissociation of percepts of motion through depth from percepts of position-in-depth using similar anticorrelated stimuli (Rokers, Cormack, & Huk, 2008a). Although it could be argued that the computations of static and dynamic disparity mechanisms may be distinct processes, we believe that the term "changing disparity" should (and has been) defined as a signal that takes conventional supra-threshold static disparity signals as its input. Thus, if it can be demonstrated that a given stimulus configuration does not support depth judgments based upon static disparities, then it cannot support 3D motion judgements based on changing disparities. This disparity hierarchy appears to hold in both classical and recent models of CD mechanisms (Cumming, 1995; Peng & Shi, 2010).

We ruled out another concern regarding the IOVD stimuli, which is that observers might have performed the direction-discrimination task on the basis of monocular direction discrimination instead of based on perceived 3D direction per se. If this were the case, the intrinsic monocular similarities of the FULL and IOVD stimuli could account for the similar sensitivities observed. Of course, this argument assumes that subjects were able to perform the task using concurrent utrocular identification and 2D direction discrimination under conditions of simultaneous stimulation in the two eyes, (which is rather unlikely, see (Ono & Barbeito, 1985; Porac & Coren, 1986), and also correctly mapping the monocular motion to the 3D direction in the absence of feedback. Regardless, we addressed this concern empirically, by performing an additional experiment in which we measured observer's frontoparallel (2D) direction discrimination sensitivity for FULL and IOVD stimuli, and compared them to their corresponding 3D motion sensitivities from the main experiment. This 2D direction-discrimination task was performed on identical stimuli as in the position-in-depth task (described above), except that observers were instructed to respond to the 2D direction of motion (leftward or rightward). The scatterplot of these data (figure 2.8, panel B) shows that sensitivity for 2D direction discrimination was several times higher than equivalent 3D motion sensitivities across all eccentricity and motion cue conditions. This replicates stereomotion suppression (Tyler, 1971) and supports a larger body of research which has shown that IOVD performance cannot be explained on the basis of monocular stimulation (e.g. Brooks & Stone, 2006a; J. M. Harris & Watamaniuk, 1995; Rokers, Cormack, & Huk, 2008a; Shioiri et al., 2000).

Although prior work has shown that the CD cue can be experimentally isolated and is sufficient to yield percepts of 3D motion (Cumming & Parker, 1994; Gray & Regan, 1996; Julesz, 1971; Norcia & Tyler, 1984), our results suggest that the IOVD cue can be similarly studied in near-isolation by using binocularly anticorrelated elements outside the fovea, moving at moderately fast speeds, at relatively sparse densities (more akin to displays traditionally used to study frontoparallel motion than those used to study stereopsis). Our results also suggest that the IOVD cue is not only sufficient to yield

percepts of 3D motion, but is also relied upon preferentially (relative to the CD cue) under many viewing conditions.

Effects of speed on the CD and IOVD mechanisms

The CD and IOVD mechanisms were affected very differently by manipulations of speed. CD sensitivity fell quickly with increased speed, exhibiting a lowpass sensitivity (or having a peak at or lower than the lowest speeds measured). IOVD sensitivity, on the other hand, had a clearly bandpass sensitivity peaking at a faster speed. FULL sensitivity also exhibited a bandpass speed tuning, one that was nearly identical to the IOVD pattern, but that contrasted sharply with that seen for the CD stimuli.

We described our stimulus motions in terms of retinal speed per eye (or, in the case of the CD stimulus, equivalent retinal speed per eye). For example, a speed of "1.8°/sec·eye" corresponds to rightward (or leftward) monocular motion in one eye at 1.8°/sec, and leftward (or rightward) monocular motion in the other eye at 1.8°/sec. Such a speed, although relatively slow when viewed monocularly, yields a percept of relatively fast motion through depth towards or away from the observer. This speed was closest to the peak of the IOVD and FULL speed tuning curves. Likewise, at 0.3°/sec·eye, the qualitative percept was of very slow displacement over time, and observers reported that this condition did not yield a "direct" perception of motion through depth— rather, the phenomenology was of inferring motion from a change in position-in-depth over time. It is therefore noteworthy that this speed yielded the highest CD sensitivity.

The effects of speed suggest that the perception of 3D motion in natural stimuli (which inherently contain both binocular cues) appears to be supported by the IOVD cue more than the CD cue. That said, the CD cue appears well-suited to carry information for very slow 3D motions. One possibility is that the CD mechanism is not fundamentally a motion mechanism, but rather one optimized for objects that are (nearly) stationary. The (isolated) CD signal can drive vergence eye movements, but these are relatively sluggish —generally less than 1°/sec (Stevenson et al., 1994). We speculate that the CD mechanism may simply reflect the brain's attempt at inferring the pattern and rate of

change of signals from well-characterized "static" disparity detectors. In contrast, the IOVD mechanism appears to be more similar to other (2D) motion mechanisms, and is well-suited to quickly moving objects. The pattern of sensitivity we observed to the FULL cue displays further suggests that the human visual system is capable of relying on this IOVD mechanism. This is ecologically appropriate, given that sensitivity to objects moving quickly towards or away (even if they have not yet been fixated) is likely a major element in successful interaction with a dynamic 3D environment.

Eccentricity effects on the CD and IOVD mechanisms

The patterns of sensitivity to the CD and IOVD cues across a wide eccentricity range complement the observed effects of speed. Sensitivity to FULL, CD, and IOVD stimuli decreased at larger eccentricities. When considered in isolation, the manipulation of eccentricity was actually less effective at distinguishing between the CD and IOVD cues. One might have expected that CD sensitivity would be particularly affected by eccentricity because the processing of static disparities is known to be much better in the fovea (e.g. Blakemore, 1970; Tyler, 1975). However, displays that simulate reasonable 3D motions subtend a disparity range that is at least an order of magnitude greater than the stereoacuity threshold at the eccentricities tested (Westheimer & Truong, 1988). Thus, even the most eccentric stimuli (11°–15°) still contained range of disparities that could likely have supported CD sensitivity at higher eccentricities.

Instead, the effects of eccentricity are more informative when one considers them in conjunction with the manipulation of speed. IOVD sensitivity was relatively more robust to eccentricity for faster speeds of motion through depth. This improvement with speed is further consistent with the notion that our anticorrelated IOVD stimulus tapped a motion mechanism. For example, sensitivity to temporal frequency is if anything improved at far eccentricities (Rovamo & Raninen, 1984; Wright, 1987). Indeed, the eccentricity effects on IOVD sensitivity were lowest at fast speeds, perhaps because performance had achieved a maximal level. In contrast, eccentricity effects on CD

sensitivity were lowest at fast speeds as well, but in this case, the reason was because performance was approaching chance (instead of peak performance) levels.

Taken together, the consideration of speed and eccentricity effects suggest that the CD mechanism is capable of supporting 3D motion direction discrimination for slow, (and particularly para-foveal) motions. Outside of this range, the IOVD mechanism appears much more capable of accounting for 3D direction discrimination when both cues are present. We interpret this as evidence for the relative primacy of the IOVD cue outside central vision. The visual system may exploit interocular velocity differences as a robust source of information for moderate- and fast- moving objects that one is not (yet) looking at. This suggests that classical motion detectors, typically studied in the domain of 2D processing, may also be utilized for perceiving 3D motion. The question as to how neural circuits implement the differencing operation upon eye-specific velocity signals remains an open question and a topic of ongoing work.

Relation to past work

Although our results demonstrate that the IOVD cue makes a significant contribution to 3D motion perception that is sometimes superior to the CD cue, it is important to recognize that this role for IOVDs may depend to some extent on the experimental conditions (Regan & Gray, 2009). For example, we used a 2-alternative forced choice direction-discrimination task. We selected this task because it seemed most analogous to a particularly well-studied task in the 2D motion literature (e.g. Newsome & Paré, 1988; Watamaniuk, McKee, & Grzywacz, 1995), but prior work has (understandably) investigated 3D motion perception using a variety of different tasks, including direction estimation, speed discrimination, judging time to contact, and indicating whether motion through depth is perceived (Brooks & Stone, 2006a; J. M. Harris & Dean, 2003; J. M. Harris & Watamaniuk, 1995; Portfors-Yeomans & Regan, 1996). Because each of these tasks might require the observer to rely on and interpret 3D motion signals in different ways—ways that we do not yet fully understand—it is difficult to generalize or compare results across tasks. Indeed, it will be interesting to

extend the approach described in this paper to these other tasks in order to build a broader characterization of the relative contributions of the CD and IOVD cues. Although many of these tasks tap important perceptual capacities, we again emphasize that our conclusion that the velocity-based cue plays a major role may reflect the fact that the task and stimuli we chose had strong roots in literature on both the psychophysics and physiology of 2D motion processing (e.g. Braddick, 1974; Newsome & Paré, 1988; Perrone & Thiele, 2002). Given this constraint, it also had an obvious real-world validity (judging whether something is moving towards or away from your head).

Driven by the goal of maintaining consistency across FULL, CD, and IOVD conditions, we decided to use a single signal plane, which was easily depicted in all cue conditions, salient at high signal strengths, and allowed for straightforward manipulations of both eccentricity and speed. However, it will be important to generalize these results to other stimulus conditions in order to relate it to a larger body of prior work. Some prior work has employed small stimuli relatively near fixation, various types of spatial motion structure (i.e. sinusoidal oscillation, rotation, or oblique trajectories through depth), and element densities ranging from a few percent to complete coverage— all of which could affect the relative contributions of the two cues (Andrews et al., 2001; Cumming & Parker, 1994; Portfors-Yeomans & Regan, 1996; Shioiri et al., 2008). Of course this dependence on specific stimulus factors is true whenever one studies a system that can use multiple sources of information. The key point here is that we have found a set of reasonable and simple conditions under which the IOVD cue makes a surprisingly strong contribution to the perception of 3D motion.

We used stimuli that moved across a wide range of constant speeds consistent with motions of real objects through depth: at our viewing distances, we simulated 3D motions of about 8 cm/sec to about 74 cm/sec (the latter corresponds to approximately one and three-quarters miles per hour; a reasonable walking speed for a human). Across this range, we observed large (approximately an order-of-magnitude) changes in overall sensitivity, as well as large relative changes between the CD and IOVD conditions. These

changes in sensitivity suggest that the IOVD cue makes a major contribution to 3D motion perception at ecologically-important speeds.

An important prior study concluded that IOVDs did not contribute to 3D motion perception across a very wide range of temporal frequencies (Cumming & Parker, 1994). There are several reasons that might explain why we arrived at a starkly complementary conclusion. First, we directly assessed the IOVD contributions using anticorrelated displays, instead of inferring them from the difference between FULL-cue and CD-only displays (although note that, under our experimental conditions, the latter method would still have revealed a large role for IOVDs, as shown in Figure 2.8). More importantly, we asked observers to perform a single interval direction-discrimination task on stimuli moving at a constant speed, which is quite different from the prior study's use of a two interval signal-present vs. signal-absent task on stimuli that oscillated sinusoidally through depth. It is possible that subjects could identify the presence of the signal in this discrimination task by preferentially attending to the slower parts of the sinusoidal oscillation at the extremes of the depth range. Such a strategy could be supported almost exclusively by disparity-based mechanisms, and would reveal little about the sensitivity of IOVD vs. CD mechanisms.

Despite some significant differences between our experiments and prior ones that arrived at different conclusions, our finding of a central role for IOVDs does not imply that our CD stimulus was somehow weak or at a particular disadvantage relative to the other stimuli in our study. In fact, we found that sensitivity in the CD condition actually exceeded that in the FULL (and IOVD) conditions at the slowest speeds, demonstrating that the CD stimulus itself contained strong signals under the viewing conditions that favored CD processing—thus, the relative inability of the CD condition to account for FULL sensitivity at faster speeds almost certainly lies within the visual system. The position-in-depth control experiment lends further support: for the same signal and noise dots, performance on a position-in-depth task was nearly perfect for the CD stimuli, but abysmal for the IOVD stimuli. Although not conclusive, this is certainly evidence against the notion that CD signals were at a huge disadvantage due to low level masking.

Moreover, the upper limit of speed sensitivity that we observed is not at odds with some prior studies, most notably that of Cumming and Parker (1994). They collected data at slower speeds, but the sensitivity in their temporally-correlated condition (analogous to our FULL condition) clearly improves with temporal frequency for both subjects shown, while the sensitivity in their dynamic condition (analogous to our CD condition) suggests a roll-off around 2Hz. Although both of their subjects were, in fact, better overall in their "CD" condition at the temporal frequencies tested, the difference in sensitivity between the two conditions was also clearly diminishing rapidly with increasing temporal frequency: from their Figure 2.3, it is not at all unreasonable to suppose that the sensitivity in their "FULL" condition would begin to exceed that of the "CD" condition had higher temporal frequencies been tested.

In another study, Norcia and Tyler (1984) found that a CD-based depth percept was present up to 6 Hz, but they used a square-wave alternation in depth and noted that the percept changed from one of apparent motion (in depth) to one of pulsating semi-transparent depth planes as temporal frequency was increased. It is thus unclear what portion of their responses can be attributable to true 3D motion percepts, and what portion was due to a modulation of signal strength at different disparities. Moreover, as they themselves noted, their estimate of the temporal resolution of stereoscopic position change was higher than had been reported in previous work (Regan & Beverley, 1973; Richards, 1972). Overall, we find that the similarity of results to prior work (despite the differences with earlier studies' initial conclusions) yields a rather coherent picture of the relative temporal sensitivity of the CD and IOVD cues.

More recent work has also provided evidence for characteristic dependencies of the IOVD cue on speed and eccentricity. Shioiri et al. (2008) reported greater sensitivity to higher temporal frequencies for uncorrelated dot displays (presumably mediated primarily by the IOVD mechanism) than for correlated cyclopean displays (processed exclusively by the CD mechanism). Although observers performed different tasks in the IOVD and CD conditions (single interval direction discrimination versus two interval signal detection, respectively), and different motion characteristics were present in the

two conditions (rotation in depth versus oscillation in depth, respectively), the general conclusions they arrived at are rather consistent with our speed tuning observations. Likewise, Brooks & Mather (2000) reported evidence for an IOVD contribution to 3D motion based upon manipulations of eccentricity. Reductions in perceived frontoparallel speed at farther eccentricities mirrored reductions in perceived speed of 3D motion, but were relatively independent of eccentricity effects on disparity-based judgments. Such a result is consistent with the robust IOVD contributions across speeds that we observed at middle and far eccentricities.

More generally, our results complement prior attempts to isolate IOVD contributions using a variety of different approaches. Although stimuli containing only CD information without any IOVDs can be straightforwardly generated using 1-frame dot lifetimes, a stimulus containing only IOVDs without also containing any potential CD information has not been developed (and may be impossible). Thus, prior work has employed uncorrelated elements (Brooks, 2002b; Shioiri et al., 2000), vertically-unmatched strips of opposite motions (Shioiri et al., 2000), or monocular adaptation (Brooks, 2002a; Fernandez & Farell, 2006). Although details of each of these approaches require careful consideration (e.g., ruling out spurious disparities in uncorrelated stimuli, assessing the effects of optical blur and neural spatial summation in vertically-unmatched strip stimuli, and understanding the relationship between monocular adaptation and subsequent dichoptic 3D processing), many of these studies have included careful controls and have begun to form a coherent and compelling case for the importance of IOVDs in 3D motion perception. The overall body of relevant work, including ours, thus encompasses a wide range of tasks and stimuli. Despite this heterogeneity, there is broad agreement that IOVDs do make a distinct contribution to 3D motion perception. Furthermore, they point to the generalization that the CD mechanism is generally lowpass, even if estimates of the cut-off speed may vary slightly. Moreover, previous studies are generally consistent with the notion that the IOVD cue supports the perception of motion through depth at relatively high speeds—speeds that are beyond the upper limit of dynamic disparity processing. Our ability to compare FULL, IOVD, and CD

sensitivities using a common stimulus geometry, task, and sensitivity metric allow us to further suggest that the IOVD cue not only contributes to 3D motion perception, but is in fact dominant in a variety of important conditions.

Implications for future work

At a practical level, our results demonstrate the feasibility of studying the IOVD cue in isolation, or at least in near-isolation. The use of anticorrelated displays provides a straightforward means for degrading disparity-based signals to reveal the role of the IOVD cue, while maintaining a simple stimulus geometry that supports direct comparison to other 3D motion displays. At a theoretical level, our results provide strong motivation to extend models of motion processing to consider the interocular comparison of monocular velocities. Canonical models of motion processing typically assume that later stages of motion processing operate on generic cyclopean representations (i.e. binocular properties are left unspecified), and thus the representation of eye-specific motions has not been considered (Perrone & Thiele, 2002; Rust, Mante, Simoncelli, & Movshon, 2006; Simoncelli & Heeger, 1998). Because motion towards or away the observer typically yields opposite directions of motion in roughly corresponding parts of the two retinae, standard motion mechanisms that involve directional antagonism (motion opponency) need to be modified to be specifically monocular (Tailby, Majaj, & Movshon, 2010). Instead of subtracting these locally-opposite directions of motion (for a net result of zero), the visual system must instead extract their signed difference as a cue to 3D velocity.

Furthermore, our results also motivate extensions of models of binocular processing to consider the contributions of monocular motions. Instead of being depicted as an "impurity" relative to CD-only cyclopean stereomotion, our results support a complementary perspective: that the IOVD cue be considered an integral part of seeing motion in depth, and that, at least under a wide range of reasonable experimental conditions, the CD cue makes a rather limited contribution. The perception of 3D motion

may thus better be thought of as a binocular form of motion processing, rather than as a dynamic form of stereopsis.

Chapter 3: Neural representation and directionally-selective mechanisms of binocular 3D motion processing[4]

INTRODUCTION

There is a wealth of psychophysical and physiological evidence for the existence of neurons tuned to roughly frontoparallel (2D) directions of motion in the primate visual system (e.g. as presented on a plane perpendicular to the observer's line of sight; (Born & bradley, 2005; Burr & Thompson, 2011). In contrast, there is relatively little evidence for the existence of neurons tuned to 3D motion (e.g. toward or away from the observer; Akase, Inokawa, & Toyama, 1998; Cynader & Regan, 1982; Maunsell & Van Essen, 1983; G. F. Poggio & Talbot, 1981; Regan & Cynader, 1982; Toyama, Komatsu, Kasai, Fujii, & Umetani, 1985; Zeki, 1974). Further, prior psychophysical work has revealed that such 3D motion processing relies both on estimating changes in binocular disparity over time and by comparing different monocular velocities across the two eyes (J. M. Harris et al., 2008; Regan & Gray, 2009). It remains unclear whether such computations are explicitly represented by later processing stages that are directionally-selective for 3D motion. Indeed, given the scant electrophysiological evidence for 3D tuning in individual neurons compared to the widespread occurrence of 2D tuning throughout visual cortex, one might wonder whether such disparity- and velocity- based inferences are not explicitly represented by 3D direction-selective neural populations, but are instead extracted by cognitive and motor circuits that only "read out" 3D direction when required for task performance or action.

In the first experiment, we employed the motion aftereffect (MAE) to test the hypothesis the visual system contains neural populations tuned to 3D directions of motion. Following the prior logic of 2D MAEs, we reasoned that prolonged viewing of

---

[4] This chapter is based on:
Czuba, T. B., Rokers, B., Guillet, K., Huk, A. C., & Cormack, L. K. (2011). Three-dimensional motion aftereffects reveal distinct direction-selective mechanisms for binocular processing of motion through depth. *Journal of Vision*, 11(10), 18. doi:10.1167/11.10.18

unidirectional motion toward an observer would make subsequently viewed stimuli more likely to appear to be moving away (and vice versa). Such an aftereffect could be interpreted as the result of a post-adaptation imbalance of responses of neurons tuned to motion toward (weaker after adaptation toward) versus those tuned to motion away (unaffected by adaptation toward) (Anstis, Verstraten, & Mather, 1998; Barlow & Hill, 1963; Mather, 1980). Alternatively, the lack of a MAE would suggest that neurons tuned to 3D motion do not exist, and rather that later stages of cognitive and motor processing, which presumably do not adapt, are involved in a less explicit process of inferring 3D motion from the responses of a population of neurons that are themselves not selective for 3D motion, but code the relevant building blocks.

Although such "if you can adapt it, it's there" logic (Mollon, 1974) has repeatedly been applied to the case of 2D motion, the interpretation of 3D motion aftereffects requires additional care. One major interpretive challenge is due to the fact that 3D motion processing depends at least in part on exploiting the fact that objects moving toward or away from an observer project different horizontal directions of motion to the two eyes (Brooks & Stone, 2004; Czuba, Rokers, Huk, & Cormack, 2010; Rokers et al., 2009; Shioiri et al., 2000). Thus, to interpret a 3D MAE as unambiguous evidence for the existence of mechanisms tuned to 3D direction, one must distinguish 3D motion adaptation per se from the inherited adaptation effects of the monocular 2D mechanisms that send signals to the putative 3D mechanism. We addressed this issue by separately measuring the 2D monocular MAEs and then testing whether they could quantitatively account for the magnitude of the 3D MAE.

Such a directionally-selective representation of 3D motion could be based on binocular mechanisms specific to processing motion through depth. In the second experiment we address the two primary binocular motion cues that could be contributing to the 3D MAE: a disparity-based, changing disparity (CD) cue; and a velocity-based, interocular velocity difference (IOVD) cue.

The CD cue, changing disparity over time can be computed by taking the time derivative of horizontal binocular disparity (i.e. comparing an object's changing position-

in-depth over time; Cumming & Parker, 1994; Gray & Regan, 1996; Regan & Gray, 2009). The CD cue has traditionally received a great deal of attention because of a compelling ability to generate 3D motion percepts through purely cyclopean pathways, i.e. stimuli that are completely devoid of coherent 2D (monocular) motion signals (Julesz, 1960). This is achieved by dynamically relocating stimulus elements on a plane frontoparallel to the observer on successive display frames, while presenting a series of steadily changing binocular disparities that correspond to motion toward or away from the observer.

The IOVD cue, interocular velocity difference, takes advantage of the geometry of binocular viewing, wherein an object moving through depth creates different (and often opposite) directions of motion in the two eyes. The direction of 3D motion can therefore be computed by directly comparing monocular velocity signals in corresponding regions of the two retinae. Although the IOVD cue was first proposed by Beverley & Regan in 1973, it was not critically addressed until two decades later (Cumming & Parker, 1994; Portfors-Yeomans & Regan, 1996), and until recently, has been thought to make little or no contribution to 3D motion processing (J. M. Harris et al., 2008; Regan & Gray, 2009). There is however a growing body of evidence suggesting the IOVD cue plays an important, if not primary, role in a variety of ecologically plausible viewing conditions (see Discussion).

In the second experiment, we examined the relative contributions of the cues potentially underlying the 3D MAE. We approached this using cue-isolating adaptation stimuli to differentially measure the contributions of the velocity-based and disparity-based cues to the full 3D MAE. Further, by adapting to isolated CD or IOVD stimuli, and testing with an identical test stimulus (containing both cues), we were able to examine the relative contribution of each binocular cue to the representation of 3D directions of motion.

In summary, we performed a series of psychophysical experiments identifying a distinct 3D MAE, using methods that allowed us to consider the contributions of monocular adaptation and to quantify the relative contribution of the CD & IOVD cues.

Our results reveal a surprisingly large 3D MAE, provide psychophysical evidence for the existence of neurons tuned to 3D direction of motion, and demonstrate that the motion aftereffect can be used to probe the mechanisms of 3D motion perception.

METHODS & MATERIALS

Observers

Data were collected in three psychophysical observers (three of the authors, males aged 27–47), all with good stereopsis, and normal or corrected-to-normal vision. Experiments were undertaken with the written consent of each observer, and all procedures were approved by the UT-Austin Institutional Review Board. A total of 7776 trials were collected across the 3 observers. The nature of motion adaptation experiments required the use of highly experienced observers capable of maintaining continuous fixation for the entire duration of each experimental session (17–20 minutes at a time, 17.8 hours total).

General procedure

We measured the magnitude of motion aftereffect using a motion nulling paradigm (Blake & Hiris, 1993). Following adaptation to unidirectional 2D or 3D motion, we presented a series of test stimuli which contained variable motion coherence in the same or opposite direction of adaptation interleaved with brief top-up adaptation stimuli (see General stimuli). On each trial, observers reported the perceived direction of test stimulus motion in a 2-alternative forced choice task, responding either leftward/ rightward, or toward/away; depending on condition as appropriate. No feedback was provided. Using the method of constant stimuli, direction discrimination sensitivity was measured across a range of motion coherences by adjusting the ratio of signal dots to noise dots.

General stimuli

Observers stereoscopically viewed (via mirror stereoscope; see Apparatus & displays) moving random dot displays in which 80 dark (0.55 cd/m2) or light (124.25 cd/m2) binocularly paired dots were presented on a mid-gray (61.40 cd/m2) background. In each monocular half-image, half the dots were dark and half the dots were light (Figure 3.1, screenshot of stimulus spanning two monitors). Individual dots subtended a visual angle of 9 arcmin (0.15°) and were anti-aliased to achieve subpixel position accuracy. Stimulus dots were uniformly distributed within a volume spanning 2.5–8° in eccentricity and ± 72 arcmin disparity. Observers fixated on a single, static, bright stimulus dot in the center of a small central square (0.5°) with horizontal (black) and vertical (red) nonius lines located in the center of each monocular half-image. To further aid fixation and confirm proper binocular alignment, a static 1/f noise texture surround was presented around fixation marks (0–0.75° eccent., 0 arcmin disparity) and beyond the stimulus annulus (≥ 9.8° eccent., 0 arcmin disparity).

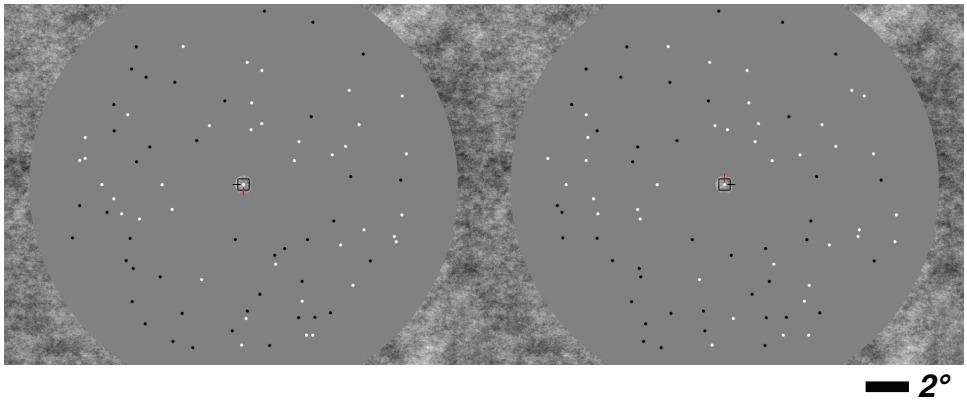

Figure 3.1 Screen capture of the basic stimulus

In the actual experiments, the right and left halves were split between two monitors and viewed through a mirror stereoscope. Nonetheless, free-fusing will give a reasonable impression of the experimental percepts. We found that the 1/f texture in the center and surround greatly facilitated fusion and holding stable vergence (which subjects could monitor with the horizontal and vertical nonius lines at fixation).

Signal dots moved at a monocular speed of 0.6°/sec. Upon reaching the edge of the stimulus volume, dots were wrapped to the opposite end of the volume and were randomly relocated to minimize apparent motion during the wrap. Thus, dot lifetimes were constrained by the duration of travel through the stimulus volume. Because this constraint was most severe in 3D motion conditions ([depth of volume] / [frame rate] = 120 frames = 2.0 s), identical adaptation dot lifetimes of 2.0 seconds were imposed for frontoparallel and monocular motion stimuli.

While monocular motion speeds of 0.6°/sec are relatively slow compared to those used in most 2D motion research, when stimuli are moving in opposite directions between the two eyes, even relatively slow monocular speeds correspond to brisk 3D motion speeds. Furthermore, at similar speeds and retinal eccentricities, direction discrimination sensitivities are approximately equivalent for the two primary binocular 3D motion cues (Czuba et al., 2010).

After an initial adaptation period (100 s), observers were presented with a series of test stimuli (1 s) moving in the same or opposite direction as adaptation followed by brief top-up adaptation stimuli (4 s; 1.25 s inter-stimulus interval) to maintain steady-state adaptation. Test stimuli were similarly distributed throughout the stimulus volume, but had brief dot lifetimes (15 frames, 250 ms), and were presented in a range of motion coherence levels (ranging from 0–95% coherence). Short dot lifetimes were selected to reduce perceptual segregation of real and illusory motion, while still providing a clear motion percept and a useful dynamic range in resulting psychometric functions (Lankheet & Palmen, 1998; Watamaniuk et al., 1995). Observers reported the perceived direction of test stimulus motion with a left or right mouse click.

Manipulation of 3D motion coherence

In the test stimulus of all experiments, 3D motion coherence, defined as the ratio of signal dots to noise dots (e.g. Newsome & Paré, 1988), was randomly varied on a trial-by-trial basis according to the method of constant stimuli. Signal dots moved coherently and uniformly in the same or opposite direction as the adapter, while noise dots moved in

random walks along that same dimension. Regardless of their signal/noise designation, all dots (excluding those exceeding the stimulus volume) were displaced 0.01° either toward or away from the observer on every display frame.

Based on pilot experiments, motion coherence levels were selected to span the dynamic range of observers' responses before and after adaptation. For 3D test conditions (3D, IOVD, CD, & 3D-planar; see Expt. 1 & 2 Methods for detailed description), direction discrimination was measured at coherence levels of ± 5, 20, 50, 80, and 95%. Figure 3.2 shows an illustrative gradient of coherence for frontoparallel (2D) motion stimuli; each coherence panel depicts a single eye's monocular half-image (in this case, the right eye). By arbitrary convention, we define leftward/away motion coherence as negative coherence, and rightward/toward motion as positive coherence. For most[5] 2D adaptation conditions, a narrower range of ± 5, 20, and 50% coherence provided sufficient coverage for convergence of psychometric fits.

[5] Sampling resolution of monocular MAEs (mono-adapt, mono-test) was increased to 5, 12.5, 20, 25, and 50% coherence.

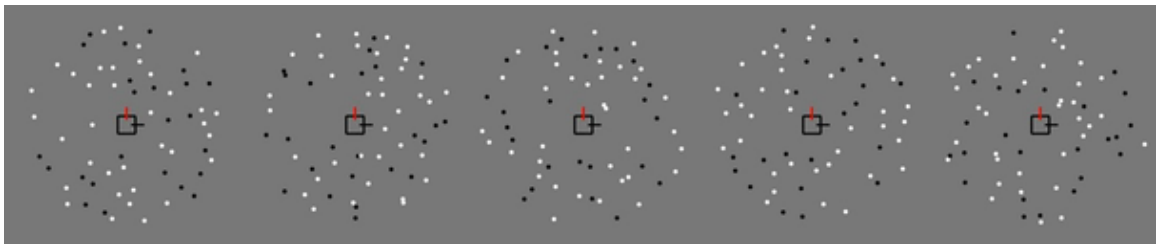

Figure 3.2 Motion coherence manipulation (Movie 1)

A (2D) sampling of various coherence levels as used in our method of constant stimuli. Each panel represents a single monocular half-image across a gradient of motion coherence levels (± 5, 20, 50, 80, and 95%). The reader should readily appreciate a continuum of motion strength.

Motion coherence was pseudo-randomized across trials within a run. Each run consisted of a single adaptation direction with 12 to 24 trials per coherence level (6 to 10 coherence levels, depending on condition), for a total of 120 to 144 trials per run. Each observer completed 2 runs per condition in randomized order, with a minimum of 30 minutes between consecutive runs. When feasible, data from compatible conditions (e.g. monocular MAE from left and right eyes) were combined across ocular pairs (Table 3.1).

Apparatus and displays

Stimuli were presented on a pair of linearized 19" CRT monitors (Viewsonic G90; 60Hz progressive scan, 1024 x 768 pixel resolution per display) viewed through a mirror stereoscope. The monitor was driven by Mac Pro computer equipped with an NVIDIA GeForce 8800 GT video card.

Monocular half-images were presented separately on the two monitors, with a septum and various baffles positioned to assure that each monitor was only visible to the corresponding eye. Viewed through the 90 cm optical path length of the stereoscope, each monocular half-image subtended 22° of visual angle. The displays were driven using a dual-monitor-spanning video splitter (Matrox DualHead2Go) to ensure frame-locked temporal synchrony between the two displays. All stimuli were generated using the Psychophysics Toolbox (Brainard, 1997) and MATLAB (2007a, The Mathworks Inc., Natick, MA).

Table 3.1. Adaptation & test motion condition matrix

MAE Condition	⟨Adapt, Test⟩	Adapt Directions	Response	# coh	Total trials	Data in Figure #
3D	⟨3D, 3D⟩	Toward | Away	Toward | Away	10	720	3.5
2D	⟨2D, 2D⟩	Left | Right	Left | Right	6	432	3.6
Monocular	⟨mono, mono $_{same}$⟩	Left | Right	Left | Right	10	1440*	3.7
3D-mono	⟨3D, mono⟩	Toward | Away	Left | Right	6	864*	3.8
Interocular transfer	⟨mono, mono $_{opp.}$⟩	Left | Right	Left | Right	6	864*	3.8
IOVD	⟨IOVD, 3D⟩	Toward | Away	Toward | Away	10	720	3.12
CD	⟨CD, 3D⟩	Toward | Away	Toward | Away	10	720	3.12
3D-planar	⟨3D, 3D⟩	Toward | Away	Toward | Away	10	720	3.12
Unadapted 3D	⟨—, 3D⟩	—	Toward | Away	6	432	3.5
Unadapted 2D	⟨—, 2D⟩	—	Left | Right	6	432	3.6
Unadapted Monocular	⟨—, mono⟩	—	Left | Right	6	432	3.7

** combined across ocular pairs*

Data analysis

Data were analyzed by computing observers' proportion of toward/rightward responses as a function of test motion coherence. For each condition, we combined data across multiple runs for each subject, and fit psychometric functions (2-parameter logistic; eq. 1) to data collected before and after adaptation:

$$f(x) = \frac{1}{1 + e^{-2\alpha(x+\beta)}} \qquad \text{(eq. 1)}$$

Because fitted logistic parameters, β (shift) and α (slope), are equivalent to the effect adaptation had on the perceived direction and sensitivity of observer's direction discriminability, all analyses were performed directly on these fitted parameters. Bootstrapped confidence intervals on the logistic parameters were computed by resampling (with replacement) the binomial responses from each subject to create 1000 repetitions of the experiment, and fitting a psychometric function to each resampled experiment. To improve estimates of the slopes of adapted psychometric functions,[6] a single slope was fit for each observer–motion condition (i.e. raw data for each adaptation direction shifted by β of an initial logistic fit, to which a single slope [α] was fit). Finally, fitted logistic parameters were averaged across individual observers to yield a single bootstrapped distribution of psychometric fits for each motion condition. Plotted psychometric functions correspond to the median fit parameters (after checking that the median values were very similar to the means). Individual data points are derived from raw data averaged across observers, and are presented to provide a sense of variability across observers. The magnitude of motion aftereffects were estimated from the motion coherence level at which observers were equally likely to report seeing leftward/toward or rightward/away motion (the point of subjective equality [PSE]). All error bars represent 95% confidence intervals on the bootstrapped distribution (corresponding to approximately ±2 SEM for distributions that are roughly Gaussian).

[6] This was desirable because of explosive slope fits in a single instance: observer TBC, rightward 2D motion adaptation; 1 in 54 total psychometric fits on the observer-motion-direction level of analysis.

EXPERIMENT 1: METHODS

2D vs. 3D MAE.

In the first experiment, we tried to determine whether the visual system explicitly represents 3D directions of motion by adapting observers to moving dot stereograms moving directly toward or away from the observer (opposite horizontal motions in the two eyes; Figure 3.3). The resulting 3D MAE was compared against 2D (frontoparallel) MAE that were induced by presenting moving dot stereograms which contained the same, rather than opposite, horizontal motion in the two eyes (Figure 3.4). This produced a percept of a 3D cloud of dots moving leftward or rightward, frontoparallel to the observer. By only manipulating the relative monocular motions in the two eyes, this method ensured that the overall dot density, distribution of disparities, and monocular stimulation was identical across the two conditions.

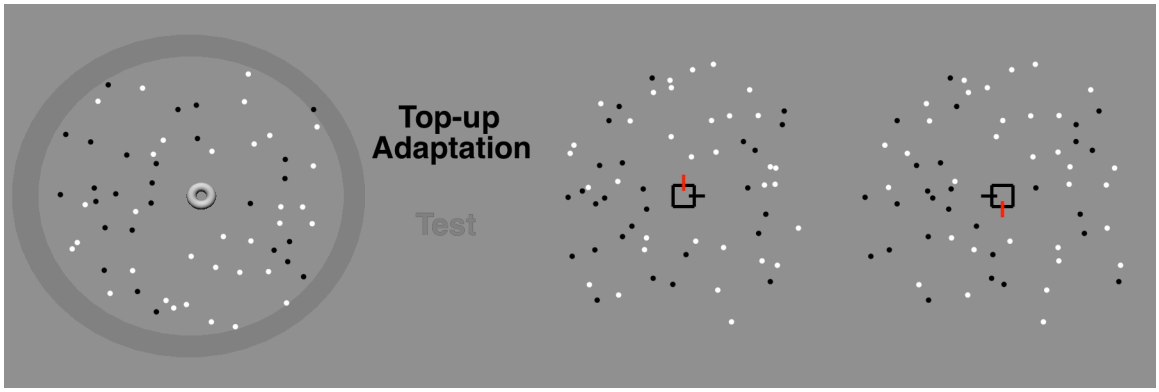

Figure 3.3 3D motion adaptation & test stimulus (Movie 2)

Depiction of the 3D motion-through-depth stimulus during a series of top-up adaptation and test presentations. The right panel shows a faithful rendition of the stereoscopic stimuli used (fusible stereo-pair), and the left panel shows a perspective view of the same stimulus sequence.

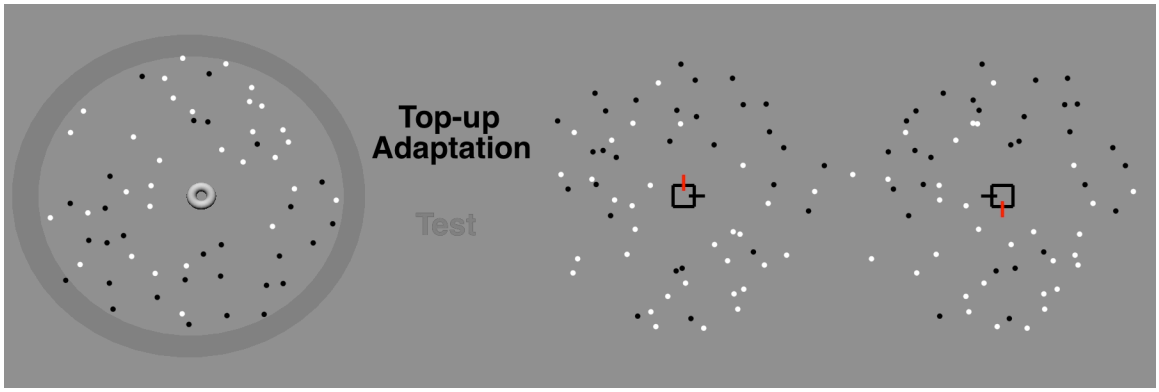

Figure 3.4 Frontoparallel adaptation & test stimulus (Movie 3)
Same as Figure 3.3, but showing the frontoparallel motion stimulus.

Dissociation of 3D MAE from inherited, monocular 2D MAEs

As noted previously, the presence of a 3D MAE is not sufficient evidence for neurons tuned to 3D directions of motion without further distinguishing it from a simple inheritance of monocular MAE. To distinguish the 3D MAE from plausible combinations of inherited aftereffects, we measured adaptation to the monocular components of the 2D and 3D motion conditions. Monocular adaptation and test stimuli were exactly the same as in the previous MAE conditions, except that one monocular half-image was replaced with a mean grey field. Having effectively removed any disparity information, the stimulus percept now appeared as a single monocularly-visible plane of dots moving leftward or rightward.

When considering whether the 3D MAE only reflects a combination of monocular aftereffects, one must also take into account the binocular interaction that occurs during adaptation to opposite monocular motion in the two eyes. We addressed this with two additional control conditions that measured: (1) the frontoparallel monocular MAE following adaptation to motion toward or away from the observer, and (2) the interocular transfer of monocular MAE.

EXPERIMENT 1: RESULTS

Existence of a 3D MAE and comparison to 2D MAE

We observed a strong 3D motion aftereffect resulting from prolonged adaptation to unidirectional motion toward or away from the observer (Figure 3.5). Panel A depicts the two adaptation conditions: binocularly presented 3D motion toward or away from the observer, followed by binocularly presented test stimuli moving in either the same or opposite direction as adaptation. Data plotted in panel B show the proportion of 'toward' responses as a function of test stimulus coherence (averaged across 3 experienced observers, 72 trials per point, 720 trials total). Increasingly away motion coherence

corresponds to negative values on the x-axis, and increasingly toward motion coherence to positive values. A value of 0.5 on the y-axis corresponds to the point at which an observer is equally likely to report either toward or away on a given trial, and indicates an observer's point of subjective equality. The black line shows a logistic fit to subjects responses prior to adaptation; essentially the observer's 3D direction discrimination sensitivity. As expected, this line is centered on zero motion coherence and has a moderate sensitivity (α-1 = 0.153, CI95 = [0.109, 0.197]). Following adaptation to toward-direction motion (green), observers were much more likely to report noisy 3D motion stimuli as moving away (β = 0.356, CI95 = [0.307, 0.403]). Similarly, after adaptation to away-direction motion (red), observers were more like to report directionally-ambiguous 3D motion as moving toward (β = -0.557, CI95 = [-0.508, -0.601]).

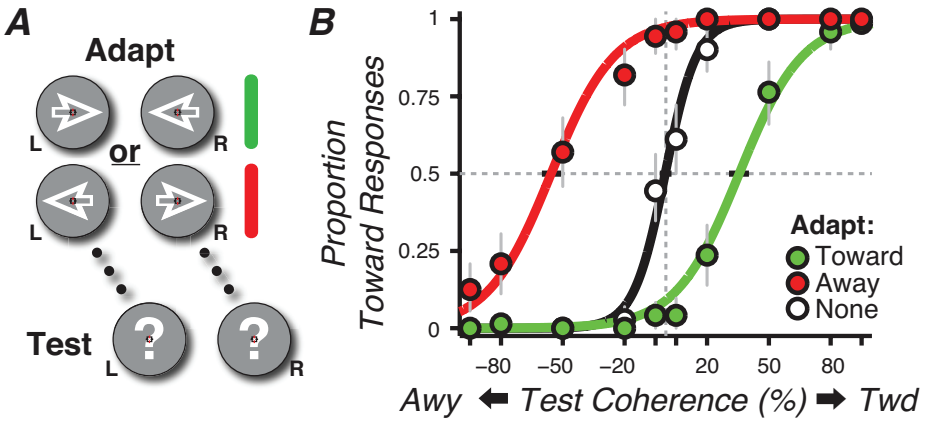

Figure 3.5 3D motion adaptation

Panel A is a schematic of the basic experimental paradigm; subjects adapted to equal and opposite motion in the two eyes (producing a 3D motion percept of dots moving either towards or away from the observer), and then judged the perceived direction of motion in depth of a test stimulus with a coherence that varied from trial to trial. Panel B shows the psychometric functions (parametrically combined across observers; see Methods) mapping the coherence of the test stimulus (x-axis) to the percent of trials judged as "toward" the observer (y-axis). The green curve is a "toward" adapter, red an "away" adaptor, and black is a reference curve collected without any adaptation. Grey error bars show bootstrapped 95% confidence intervals. The abscissa corresponding to the 0.5 ordinate on each curve represents the point of subjective equality for each condition, and bootstrapped 95% confidence intervals are shown by the black horizontal bars. Clearly a substantial 3D MAE is present.

Across a range of 3D motion coherences, these 3D MAE shifted the psychometric functions leftward or rightward (relative to an unadapted control condition), and thus could be quantified in terms of relative displacement along the x-axis, i.e., in units of the test stimulus motion coherence. 3D MAEs were equivalent to about ~45% (CI95 = [42.4, 48.9]) motion coherence. This effect struck us as surprisingly large: for the test stimuli to be judged as having no net motion on average, approximately half of the dots had to move in the direction opposite that of adaptation.

As a basis for comparison, we measured conventional 2D frontoparallel MAEs for the same observers under stereoscopic viewing conditions (Figure 3.6, panel A; dots moved in the same direction in both eyes, but were otherwise identical to those used to generate 3D MAEs). We performed the same analysis, except that the directions of motion were leftward or rightward, instead of toward or away. The y-axis now representing the proportion of rightward responses; negative values on the x-axis corresponding to increasingly leftward motion coherence, and positive values corresponding to increasingly rightward motion coherence (Figure 3.6, panel B). Unadapted direction discrimination sensitivity (black) is again centered on zero motion coherence, but with a noticeable ~2.5-fold improvement in sensitivity (α-1 = 0.054, CI95 = [0.016, 0.076]) over the 3D motion case. Increased direction discrimination sensitivity for 2D motion relative to 3D motion is unsurprising given the previously mentioned stereomotion suppression effect (Tyler, 1971). Further, Welchman, Lam, and Bülthoff (2008) have shown that greater 2D sensitivity (as measured by increment thresholds) is a consequence of a Bayesian model in which the visual system incorporates a low-speed prior; their model thus predicts this aspect of our data. What is more interesting however, is the relative magnitude of the 2D and 3D MAE. The shift in psychometric function following 2D motion adaptation was equivalent to approximately 18% motion coherence (β = 0.183, CI95 = [0.165, 0.205]). Given that previous studies on 2D MAEs using similar dynamic test stimuli have observed effects of similar magnitudes (Blake & Hiris,

1993; van Wezel & Britten, 2002), this confirms our initial impression that 3D MAEs are uniquely large.

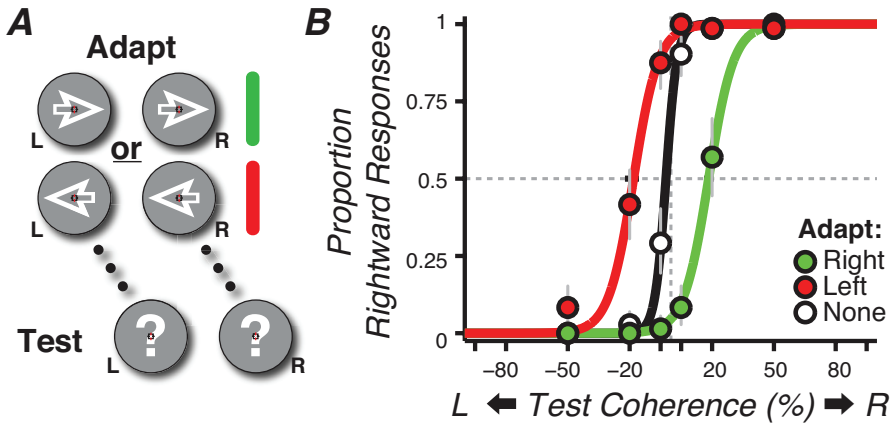

Figure 3.6 Frontoparallel motion adaptation

Schematic of (A) and data from (B) the frontoparallel motion condition. The aftereffect is much smaller than when the adaptation stimulus moved though depth, and the magnitude is also consistent with what has been reported previously for similar experiments (see text for references).

Dissociation of 3D MAE from inherited, monocular 2D MAEs

One rationale for why the 3D MAE is so large could be that it reflects multiple stages of adaptation: a 2D monocular stage (that processes the individual direction of motion for each eye, which are opposite between the two eyes), and a later 3D cyclopean stage (which extracts motion through depth after binocular integration). We therefore measured 2D monocular MAEs to assess the amount of adaptation that occurred in the earlier stages. If the magnitude of the MAEs from early stages could completely account for the 3D MAEs we observed, this would suggest that the 3D MAE was simply the result of inherited adaptation. Of course, given the fact that the 3D MAE is larger than the stereoscopically-viewed 2D MAE, this possibility struck us as extremely unlikely— but we still wished to quantify the relative contribution of the 2D stage.

We measured monocular MAEs by presenting adaptation and test stimuli in only one eye, as the observer performed the same left-right direction discrimination task as they had in the stereoscopic 2D MAE experiment (Figure 3.7, panel B). These 'pure' monocular MAEs were equivalent to approximately 20% motion coherence (Figure 3.7, panel B; $\beta = 0.198$, CI95 = [0.189, 0.208]). Although monocular MAEs were considerably smaller than the 3D MAE, their magnitudes were similar to previously reported (binocular) 2D MAE, suggesting that monocular adaptation can fully account for the 2D frontoparallel MAE.

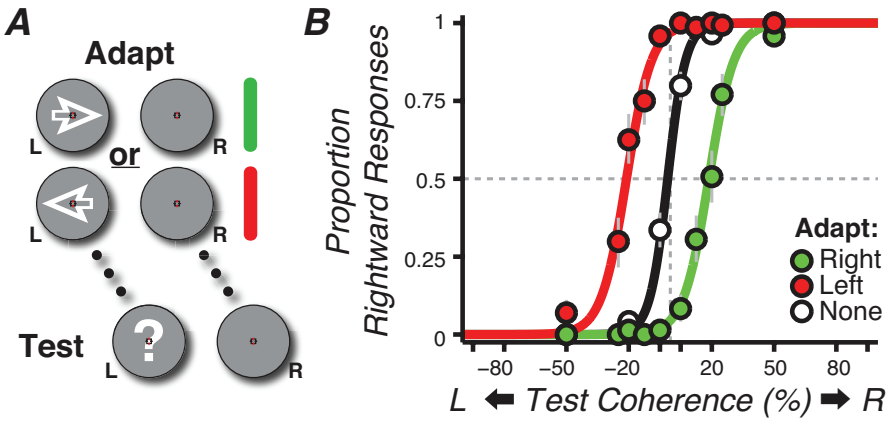

Figure 3.7 Monocular motion adaptation

Schematic of (A) and data (B) from the monocular motion condition (monocular adaptation, monocular test presented to the same eye). The magnitude of the MAE is no different (statistically) than for the frontoparallel condition shown in Figure 3.6.

If, in the 3D MAE, the two eyes' (monocular) channels adapted independently of one another, the amount of monocular adaptation in the 3D MAE could be assessed simply by adapting and testing each eye independently. However, motion processing channels exhibit significant binocular crosstalk, evidenced in varying degrees of reported interocular transfer of monocular MAE (Anstis & Duncan, 1983; Grunewald & Mingolla, 1998; Lehmkuhle & Fox, 1976). Monocular MAEs assessed after dichoptic 3D adaptation were weaker than either 2D or 3D MAEs; equivalent to approximately 9% motion coherence (Figure 3.8, panel A; $\beta = 0.093$, CI95 = [0.082, 0.104]).[7] The associated interocular transfer of monocular MAEs are shown in figure 3.8, panel B ($\beta = 0.151$, CI95 = [0.137, 0.164]).[3] The amount of monocular adaptation following dichoptic viewing of the 3D motion stimulus likely reflects a combination of direct and indirect monocular MAEs, i.e., after contamination by partial interocular transfer of the opposite direction of adaptation in the other eye. Clearly, neither scenario of monocular 2D motion adaptation could straightforwardly account for the magnitude of the 3D MAE. The relative sizes of these effects suggest that the majority of the 3D MAE arises de novo, after the adaptation of monocular mechanisms sensitive to the 2D patterns of motion falling upon each eye.

[7] In order to provide a more direct comparison between conditions plotted in Figure 8 (3D adapt-monocular test and interocular transfer) and relevant 3D motion conditions, data were combined across adaptation direction/ eye corresponding to 3D motions toward and away from the observer. Therefore, "toward" test presentations are composed of monocular test stimuli moving nasally (i.e., leftward in the right eye and rightward in the left eye), and vice versa.

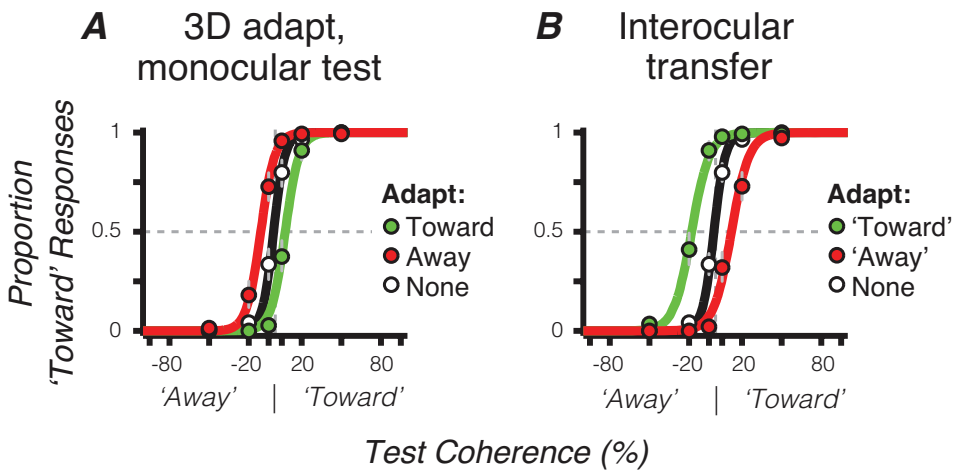

Figure 3.8 3D–mono & interocular transfer results

Data and psychometric functions from the 3D adapt, monocular test condition (3D-mono; panel A) and the interocular transfer condition (IOT; panel B). Under 3D adaptation conditions, one would expect the MAE resulting from monocular adaptation in the tested eye to be partially cancelled by the interocular transfer of the (opposing) adaptation in the untested eye. The magnitude of the 3D-mono MAE is smaller than either 2D or 3D MAEs; confirming this expectation. Panel B shows the data resulting from monocular adaptation in one eye and testing in the other eye (i.e., a direct measurement of the interocular transfer). Note that the reversal in the direction of the shift for the "towards" and "away" curves is as expected.

EXPERIMENT 2: METHODS

Binocular cue contribution

The inability to explain the 3D motion aftereffect by simple combination of monocular adaptation effects suggests that the 3D MAE must be the result of adaptation of an additional mechanism. Such a mechanism would be selective to 3D motion per se, and could compute 3D motion information based on changing disparities over time, interocular velocity differences, or both. This second experiment explores how binocular motion cues contribute to 3D adaptation by measuring adaptation to the two primary binocular 3D motion cues: changing disparity (CD) and interocular velocity differences (IOVD), using adaptation stimuli that have been shown to effectively isolate binocular motion cues. By adapting observers to stimuli that isolated either the CD or IOVD cue, and measuring the resulting MAE using a test stimulus identical to the first experiment (containing both cues), we were able to examine the relative contribution of each binocular cue adaptation to the full 3D MAE.

The general adaptation and test procedure in the second experiment was largely identical to the first experiment. After a sustained initial adaptation to unidirectional 3D motion either toward or away from the observer, subjects were presented with a series of variable motion coherence test stimuli moving in either the same or opposite direction as adaptation. The time-course of test and top-up adaptation stimuli were identical to the first experiment. The only manipulation was that adaptation stimuli were adjusted to isolate either the velocity- or disparity-based 3D motion cue.

Isolated cue conditions: CD, IOVD, and 3D-planar.

The CD cue was isolated by presenting temporally uncorrelated dots that maintain steadily changing disparities (Figure 3.9). This is achieved by relocating stimulus dots within the frontoparallel plane on every display frame (60 Hz) while presenting steadily changing disparities (Braddick, 1974; Cumming & Parker, 1994; Julesz, 1971); creating a percept similar to a plane of TV snow moving through depth.

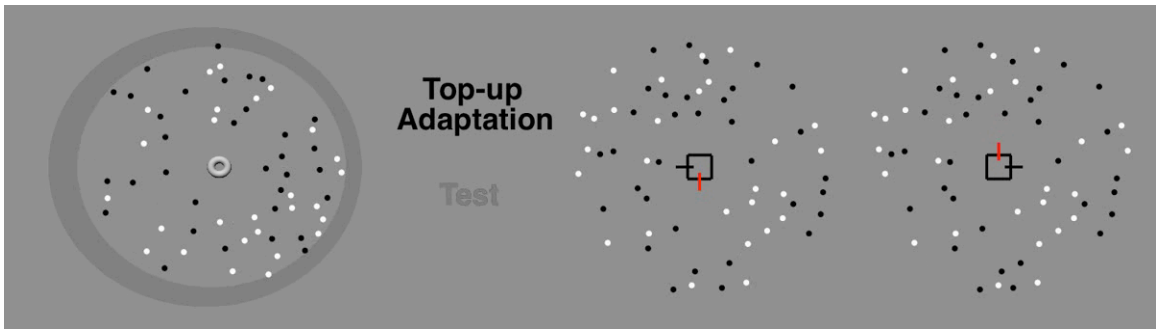

Figure 3.9 CD-isolating adaptation & test stimulus (Movie 4)
Same format as the previous stimulus movies (Figures 3.3 & 3.4), but showing the CD isolating stimulus.

The IOVD stimulus was created by pairing each bright dot in one eye with a dark dot in the other, and vice versa (i.e. anticorrelated dot contrast between the two eyes; Figure 3.10). Anticorrelation has been shown to disrupt disparity based position-in-depth information while maintaining monocular velocity information necessary for 3D motion percepts (Rokers, Cormack, & Huk, 2008a). This creates a curious percept of a stimulus that moves through depth, yet does not appear to possess a coherent position-in-depth. Rokers et al. found that the ability to discriminate position-in-depth from disparities decreased monotonically with decreasing contrast correlation, we used full anticorrelation in order to maximally disrupt the perception of depth from disparities. Nevertheless (and unlike the CD cue), interocular velocity differences cannot be geometrically isolated from disparities. Therefore, anticorrelated displays cannot be said to truly 'isolate' the IOVD cue, but to strongly bias observers towards using the IOVD cue. For the remainder of the paper, we will simply refer to the CD and IOVD stimuli with the "isolated" and "biased" qualifiers understood.

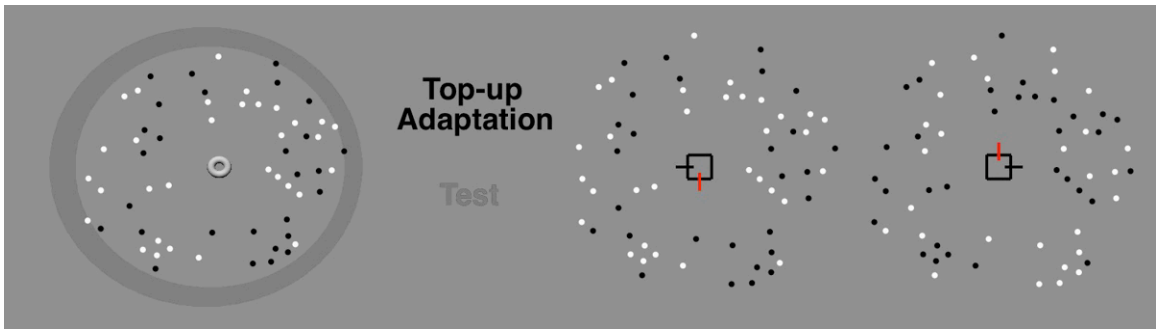

Figure 3.10 IOVD-isolating adaptation & test stimulus (Movie 5)
Same as the previous figure, but showing the anticorrelated stimulus we used to bias the observers towards using the IOVD cue. Note the relative contrast of dots presented to the left and right eye in the fusible stereo-pair (right side).

Because our disparity isolating stimulus requires a planar configuration, all cue isolating adaptation stimuli were subsequently arranged in a frontoparallel planar annulus around fixation (2.5–8° eccent.). This planar stimulus geometry was designed to provide the smoothest, most continuous 3D motion by: (1) splitting the stimulus plane into a 4-quadrant planar annulus around fixation (centered on 45, 135, 225, & 315°) with alternating quadrants distributed in depth; (2) restricting the stimulus volume to ± 54 arcmin disparity; and (3) introducing a temporal contrast ramp to the near and far limits of the stimulus volume. Quadrant pairs were distributed evenly in depth (27 arcmin, or 50% of stimulus volume, between pairs) so that at any given moment 50% of the stimulus area contained a full contrast stimulus moving in the direction of adaptation. Relative disparity of the pinwheel pairs (e.g. quadrants 1 & 3, and 2 &4) were randomized on each adaptation stimulus presentation to avoid monocular cues to 3D motion direction. Each quadrant pair progressed through the entire volume at a monocular velocity (or disparity equivalent) of 0.6°/sec, wrapping to the opposite side after reaching the depth limit. The stimulus volume spanned ± 54 arcmin (0.9°) of disparity in order to maximize the duty cycle within Panum's fusion area and avoid diplopia. A temporal contrast ramp (1/4 wave sinusoid) was applied to the outer 18 arcmin of near and far depth limits to soften the apparent motion due to the stimulus wrap. This generated percepts of alternating pinwheels of stimulus dots moving continuously toward or away from the observer. To insure that this stimulus geometry did not have an unintended effect on motion adaptation, we also measured a matched planar-wedge version of the original 3D motion stimulus that contained both binocular motion cues (Figure 3.11). To avoid confusion with the first experiment, we refer to this as the 3D-planar condition. However, the only distinction between this condition and the 3D MAE measured in the first experiment is the planar geometry; both conditions contain the same IOVD and CD cues present in naturally occurring stimuli.

For all conditions, test stimuli were identical to Experiment 1, consisting of a 3D cloud of dots randomly located throughout the stimulus volume, that contained both

binocular motion cues. This allowed direct comparison of MAE magnitude across adaptation conditions and experiments.

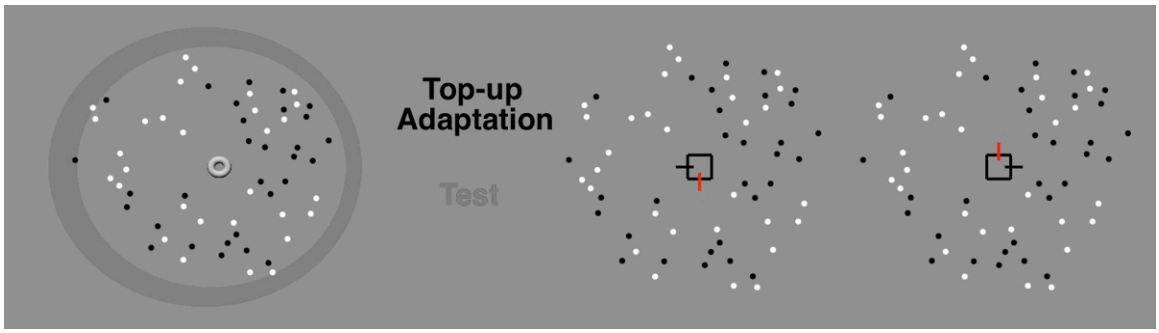

Figure 3.11 3D-planar adaptation & test stimulus (Movie 6)
Same as previous, but showing the planar wedge stimulus configuration used in Experiment 2.

EXPERIMENT 2: RESULTS

3D MAEs from isolated binocular cues

Adaptation to the planar stimulus containing both binocular cues (3D-planar; Figure 3.12, panel A) generated 3D MAE magnitudes of approximately 44% motion coherence (β = 0.440, CI95 = [0.405, 0.474]); nearly identical the original 3D MAE, confirming that the change in stimulus geometry had no effect on adaptation. Isolated velocity cue adaptation (IOVD; Figure 3.12, panel B) yielded large 3D motion aftereffects equivalent to approximately 41% motion coherence (β = 0.413, CI95 = [0.377, 0.448]), very similar to the magnitude of the 3D MAE. On the other hand, disparity cue adaptation (CD; Figure 3.12, panel C) produced markedly weaker MAE, equivalent to only 19% motion coherence (β = 0.191, CI95 = [0.164, 0.219]). Although the magnitude of the CD MAE is similar to the previous 2D MAE, the explicit lack of coherent monocular motion in our CD stimuli make it unlikely that they share an underlying locus of adaptation.

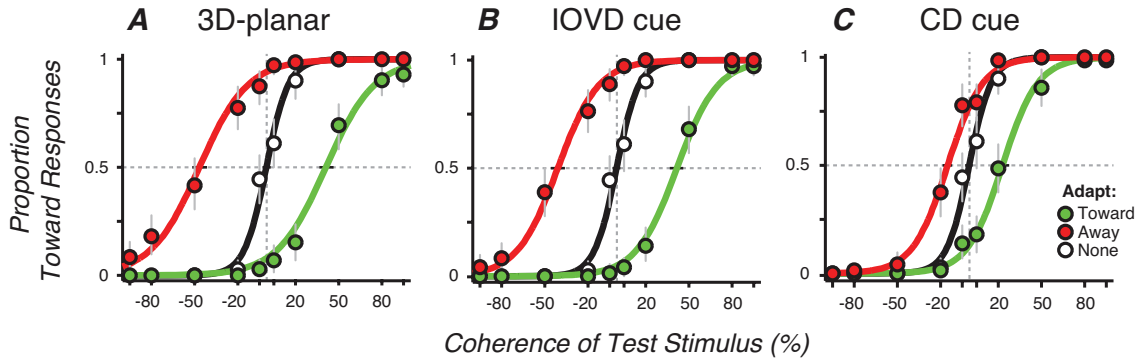

Figure 3.12 Isolated-cue adaptation results

Panel A is essentially a replication of the main data from Experiment 1 (i.e. Figure 3.5, right panel). The close agreement between the experiments indicates that specific geometry of the stimulus was of little import. Panel B shows the data from the IOVD biased adaptation stimulus and, crucially, it is nearly identical to the standard 3D MAE. Panel C shows the data from the CD isolating adaptation stimulus. Despite generation a clear depth percept during adaptation, this stimulus produced a surprisingly weak MAE.

The magnitudes of the isolated velocity-cue and disparity-cue MAEs suggest two main conclusions. First, the similarity of the velocity-based MAE to the full-cue 3D MAE implies that the 3D MAE could be fully accounted for by adaptation of a velocity-based 3D motion mechanism, without the need to consider a disparity-based contribution. Second, the mere existence of a significant (albeit smaller than the others) isolated disparity-cue MAE demonstrates that a 3D MAE can be generated using stimuli that do not contain coherent monocular motions, but which do define cyclopean stereomotion.

The overall pattern of MAE magnitudes is not only present in the combined data, but also apparent on the individual observer level. In the upper row of figure 3.13, bar graphs of MAE magnitudes for each individual observer (first three columns) reveal moderate inter-observer variability in MAE magnitudes, though the relative pattern of results across conditions seen in the combined data (fourth column) is still clearly evident on the individual observer level. The relative pattern of results is further evident when normalizing MAE magnitudes to each observer's 3D MAE (Figure 3.13, bottom row). The magnitude of MAEs generated by stimuli which contain interocular velocity cues (both 3D MAE conditions [red & orange], and IOVD condition [grey]) are distinct from all other conditions which do not contain IOVD cues.

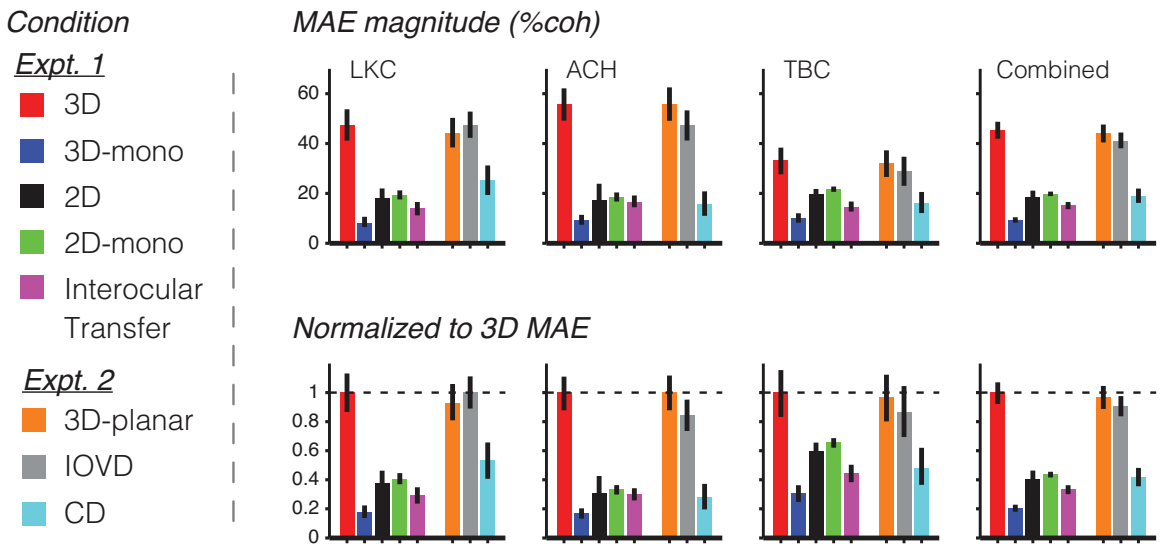

Figure 3.13 MAE magnitudes across observers & conditions

Bar graphs depicting MAE magnitudes from individual observers (first 3 columns) for all 8 motion conditions as well as the the combined data shown in previous figures (last column). The first row shows MAE magnitudes with bootstrapped 95% confidence intervals. The second row shows the same data normalized to each observer's 3D MAE magnitude.

GENERAL DISCUSSION

In summary, we observed a 3D motion aftereffect following adaptation to 3D motion toward or away from the observer that was substantially larger than the corresponding 2D MAE. We isolated the effects of monocular adaptation and interocular transfer, and found that these could not account for the magnitude of the 3D MAE. This implies the existence of a 3D motion stage that is itself direction-selective.

We then separated and isolated the two primary binocular cues to 3D motion (the disparity-based cue, changing disparity over time; and the velocity-based cue, interocular velocity differences) and compared the magnitudes of these cue-isolated 3D MAEs to a standard 3D MAE elicited by stimuli that contained both cues in concert. We found that the velocity-based 3D MAE was as large as the standard 3D MAE, confirming the central role of the velocity-based cue in 3D motion processing. We also observed a smaller disparity-based 3D MAE, demonstrating that direction-selective mechanisms can be engaged (and adapted) by stimuli that do not contain coherent monocular motions but which do specify cyclopean stereomotion.

Our two sets of experiments demonstrated that (1) the 3D MAE cannot be accounted for solely by positing "inherited" adaptation effects from earlier, monocular (2D) direction-selective mechanisms; and (2) both the velocity-based and the disparity-based binocular cues can generate 3D MAEs, although the velocity-based MAE is approximately twice as large as the disparity-based MAE, and is more similar —almost identical actually— to the standard 3D MAE that contains both cues in concert.

Quantitative comparison of all 3D and 2D MAEs

We will now analyze the results of both of our experiments together, and consider both the MAE magnitudes (i.e., shift of the psychometric functions) as well as the underlying sensitivity (i.e., slope of the psychometric functions) in each condition. This analysis further clarifies the relations and distinctions between the constituent mechanisms of 3D motion (monocular 2D direction-selectivity, interocular transfer, disparity-based "cyclopean stereomotion", and interocular velocity differences).

Figure 3.14 summarizes the results of all our experimental conditions (averaged over participants). For each condition, this plot shows the MAE magnitude (i.e. motion coherence necessary to perceptually null the MAE) on the x-axis, and the sensitivity to change in coherence (i.e., inverse slope of the psychometric function) on the y-axis. These are the two parameters from the logistic fits used to characterize observer's psychometric functions in all of the experiments. The ellipses indicate bootstrapped error ranges (Mahalanobis distances; Mahalanobis, 1936) on these two parameters (68% and 95%, corresponding to ±1 and ±2 SEMs, respectively). The plot presents these distributions of bootstrapped parameter fits for each adaptation condition (collapsed across observer and direction), and reveals more about the relations between these conditions than was clear in the separate analyses of MAE magnitudes already discussed.

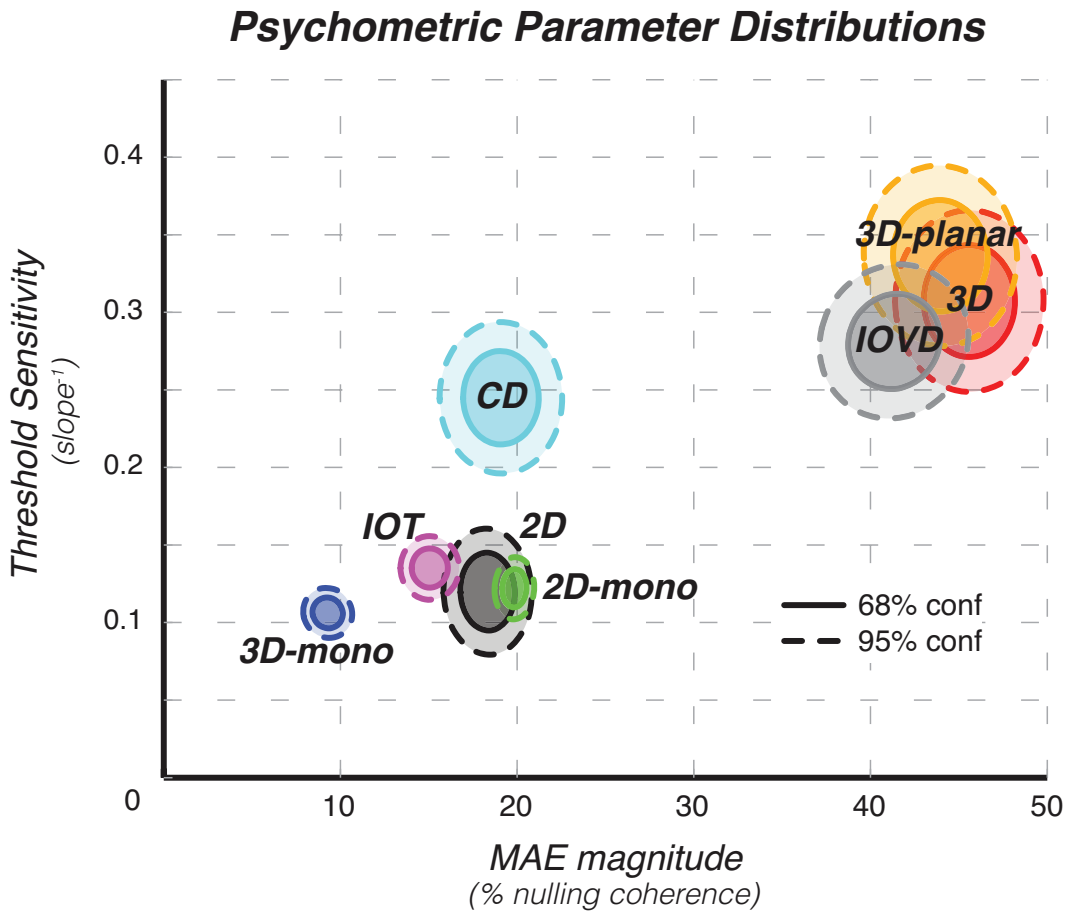

Figure 3.14 Psychometric parameter distributions across all conditions

Parametric plot summarizing the psychometric functions across all conditions in both experiments. Specifically, the steepness of the psychometric function (threshold sensitivity = α-1) is plotted on the y-axis as a function of the MAE magnitude (β) on the x-axis. The solid and dashed contours show bootstrapped 68% and 95% confidence intervals across all subjects. The most striking observation is that adaptation containing IOVDs produced similar large MAEs (3D, 3D-planar, and IOVD), while adaptation lacking IOVDs (CD and all frontoparallel conditions) yielded comparatively small MAEs.

The first thing that leaps to the eye from this parameter space plot is that the 3D MAEs (both Experiment 1 ("3D") and Experiment 2 ("3D-planar")) cluster near one another, and overlap substantially with the isolated velocity-based ("IOVD") MAE (red, orange, and grey conditions, respectively). Although our initial discussion of Experiment 1 already emphasized that the magnitudes of the 3D and IOVD MAEs were larger than those for other conditions (x-axis in this plot), it is also apparent that there is a substantial sensitivity difference (y-axis); sensitivity to 3D, 3D-planar, CD, and IOVD stimuli is lower than for the various 2D MAE conditions (e.g., "2D" [binocular], "2D-mono" [monocular], "IOT" [interocular transfer], and "3D-mono" [dichoptic 3D adaptation, monocular 2D test]; black, green, purple, and blue, respectively). This is indeed visible in the slopes of the psychometric functions previously shown (e.g., compare slopes in figures 3.5 & 3.12 to figures 3.6, 3.7, & 3.8).

Once again, both the the magnitude and sensitivity of the 3D MAE (red) is clearly distinct from the 2D MAE (black). While parameters of the dichoptic 2D MAE can be fully accounted for by monocular adaptation (2D-mono, green), monocular MAE resulting from 3D motion adaptation (3D-mono, blue) cannot account for either magnitude or sensitivity of the full 3D MAE. In fact, it is clear that both sensitivity and magnitude of the 3D MAE (in either "3D" [red], or "3D-planar" [orange]) can be fully explained by isolated velocity cue adaptation (IOVD, grey). Adaptation to the disparity isolating stimulus (CD, cyan) produces similar direction discrimination sensitivity to other 3D MAE, but with a much smaller MAE magnitude, on the order of 2D MAE. Such a sensitivity difference is broadly consistent with the phenomenon of stereomotion suppression (Tyler, 1971). The phenomenon of "two eyes being less sensitive than one" has been explained as the result of interocular motion averaging before the 3D motion computation (J. M. Harris & Rushton, 2003), which results in eye-specific motion signals that are somewhat reduced in amplitude relative to noise levels. It has also been hypothesized as the result of a bias or prior in the visual system for low retinal velocities. Welchman, Lam, and Bulthoff (2008) have shown that such a prior affects z-axis motion more, resulting in a relative decrease in sensitivity for 3D motion.

The location of the CD MAE in this parameter space plot is also informative. Again, as previously discussed, the CD MAE is smaller than the 3D and IOVD MAEs. But the existence of any MAE from CD-isolating adaptation is noteworthy, because it implies the existence of a 3D direction-selective mechanism that can be driven (and adapted) by stimuli that themselves contain no monocular motions. In the 2D motion literature, the ability of stimuli which do not contain retinal motion signals to create motion aftereffects has also been shown to occur through dichoptic combination (Carney & Shadlen, 1993) and even the mere implication of motion (Winawer, Huk, & Boroditsky, 2008). Such adaptation effects are more often taken to reflect indirect stimulation of existing motion processing mechanisms, rather than evidence for independent mechanisms specialized for each stimulus case. More directly, Patterson et al. (1994; see also Bowd, Donnelly, Shorter, & Patterson, 2000) have shown that frontoparallel MAEs induced with disparity-isolating (cyclopean) and luminance stimuli exhibit cross-domain adaptation effects. This raises the question of whether the disparity-based 3D MAE might more accurately be thought of as indirectly stimulating an underlying velocity-based 3D motion mechanism, rather than uniquely contributing to the representation of 3D directions of motion.

A subtler point that can also be gleaned from this plot is that sensitivity (slope) to the test stimulus (which contained both binocular cues to 3D motion) was similar for all 3D MAE conditions (3D, 3D-planar, IOVD, & CD), regardless of their binocular cue content. The fact that the CD MAE exhibits the same sensitivity yet lacks the distinctive magnitude of other 3D MAE further rules out the notion that the 3D MAE is simply larger due to a decreased 3D motion sensitivity relative to 2D motion. The tension between these two conclusions (i.e., the velocity-cue can fully account for the 3D MAE, but the disparity-cue in isolation can also generate a 3D MAE) is intriguing and should motivate further work about the nature of the mechanisms that combine the two binocular cues to 3D motion.

The 2D conditions also follow an interpretable pattern in this parameter space. Note that they all fall at about the same sensitivity level, regardless of whether they

involved binocular or monocular 2D test stimuli. Thus, the ordering of MAE magnitudes supports a simple set of interpretations. When viewing 2D motion in one or both eyes (and then testing in the same or both eyes), MAEs are similar regardless of whether one or both eyes are tested ("2D mono" and "2D"). This is consistent with a mechanism that is effectively cyclopean; i.e., past the point of binocular combination, so that it can be effectively and similarly driven by either eye, or both. But when one eye is adapted to 2D motion and test stimuli are presented to the other eye ("IOT", standard interocular transfer), the MAE magnitude was slightly smaller, revealing some degree of monocularity of the 2D direction-selective mechanisms.

Such IOT results have been previously reported (e.g. Mitchell, Reardon, & Muir, 1975; Raymond, 1993; Tao, Lankheet, van de Grind, & van Wezel, 2003), but the more interesting result is what happens when the observer adapted to dichoptic 3D motion and then viewed a monocular 2D test stimulus. MAE magnitude was smaller than for the other 2D conditions, but can be explained by considering the 2D-mono and IOT conditions. For each eye, there was same-eye adaptation in one direction, and (because the adapter was standard dichoptic 3D motion) there was also other-eye adaptation in the opposite direction. Thus, one might expect the resulting MAE to reflect the same-eye adaptation minus the IOT, and thus approximately equal to the 2D-mono MAE magnitude minus the IOT MAE magnitude, which seems to be roughly the case.

These results help to unpack the hierarchy of motion processing stages before and after the point of binocular combination or comparison (i.e., monocularly-biased and purely-cyclopean stages). In addition to demonstrating that a fundamentally cyclopean direction-selective processing stage exists (i.e., which cannot be accounted for by the inputs of monocular stages, and which can be adapted using purely cyclopean stereomotion), the interocular transfer results also invite speculation about the functional importance of interactions between the two monocular motion circuits. Although interocular transfer is typically viewed as an experimental convenience for assessing the degree of binocularity of a processing stage, our results also suggest that crosstalk between the left and right eyes' monocular motion stages may have important

consequences for the perception of dynamic 3D scenes. Under many conditions, when one views 3D motion, opposite directions of horizontal motion are projected onto corresponding locations in the two retinae. Because there is partial (≥50%) interocular transfer, this effectively means that monocularly-biased 2D mechanisms are not adapted as strongly by 3D motion than they would be by monocular or binocular viewing of frontoparallel (2D) motion. It is tempting to speculate that interocular transfer therefore serves a computational purpose akin to conventional motion opponency, suppressing the 2D mechanisms while still allowing 3D mechanisms to be strongly engaged. Future work is of course required to test this proposed antagonism between the two stages, and we expect 3D MAEs to be a useful psychophysical tool in this pursuit.

Relation to past work

Our results also speak to the relation between the two primary binocular cues, IOVD and CD. As we observed in a previous psychophysical study measuring 3D motion sensitivity across a wide range of speeds and eccentricities (Czuba et al., 2010), the IOVD cue is sufficient to explain the pattern of direction discrimination sensitivity when both cues are present. However, our MAE studies also demonstrate that the CD cue, although far weaker in our experimental conditions, is also a directional signal of some sort. Although a previous fMRI study by our lab implicated human MT+ as responsive to both the CD and IOVD cues (Rokers et al., 2009), cross-cue adaptation experiments should more directly assess whether the two binocular cues are integrated by a common 3D motion mechanism.

Other lines of work have investigated related, but possibly distinct, motion mechanisms. Patterson and colleagues have elicited frontoparallel MAEs using stimuli that contain the motion of disparity-defined patterns (Patterson et al., 1994; Patterson, Bowd, Phinney, Fox, & Lehmkuhle, 1996). It is not clear whether such stereoscopic / cyclopean frontoparallel motions are processed by the same changing-disparity or 3D motion mechanisms we have studied. We can only conclude that it now seems both important and experimentally feasible to begin studying the relations between 3D motion

(toward/away), 2D motion (frontoparallel), and binocular combination into cyclopean signals, despite the current convention of studying these mechanisms in isolation. Additionally, work on optic flow have demonstrated MAEs and vestibular modulations of these effects (Bunday & Bronstein, 2008; L. R. Harris, Morgan, & Still, 1981). It is also not yet clear how optic flow (classically defined as the full-field 2D pattern of retinal velocities specifying observer motion) relates to 3D motion (in our experiments, typically implemented with stereoscopic information in more localized portions of the visual field). Future investigations of the spatial scale of MAEs, and the relative contributions of monocular and binocular information to flow and 3D MAEs, may be helpful in understanding whether distinct 3D (object) motion and optic flow mechanisms exist.

Finally, some other prior work has investigated 3D motion adaptation under conditions more similar to our study. In fact, some of the earliest explorations into the presence of multiple binocular 3D motion cues (Beverley & Regan, 1973) were based on the effects 3D adaptation had on motion detection thresholds. More recently, monocular and binocular motion adaptation paradigms have been used to assess the contribution of interocular velocity signals to 3D motion processing (Brooks, 2002a; 2002b; Fernandez & Farell, 2005; 2006; Shioiri et al., 2009). Preliminary results from Sakano et al. (conference abstracts: Sakano, Allison, & Howard, 2005; Sakano, Allison, Howard, & Sadr, 2006) also reported 3D motion aftereffects from binocularly paired and unpaired (unmatched between the two eyes) random dot stimuli. In accordance with our diminished disparity-based MAE, they also reported a compelling lack of 3D MAE following adaptation to disparity-isolating 3D motion stimuli. Our results complement other work on the topic by dissecting monocular and binocular contributions to the representation of 3D motion, and we expect direction-selective adaptation and MAE experiments to enjoy a fruitful extension from 2D to 3D applications.

Conclusions

The 3D MAE provides compelling evidence for distinct neural representation of 3D directions of motion. Our assessment of binocular cue contributions show that

adaptation of the velocity-based (IOVD) mechanism alone generates a concomitantly large 3D MAE, capable of fully accounting for aftereffects generated under normal conditions, in which both IOVD and CD cues are present. These results paint an interesting picture of 3D motion processing in which the visual system explicitly represents 3D directions of motion, distinct from 2D monocular motion components, yet does so primarily based on a mechanism that compares monocular velocity signals. These MAE experiments also provide a basic experimental framework for further study of 3D motion mechanisms. Just as the MAE has been used to investigate 2D motion processing, future work can investigate factors like the spatial and temporal tuning of adaptation to further characterize this mechanism. Ideally such studies can dovetail with similar adaptation protocols in human fMRI experiments (Rokers et al., 2009), & (conference abstract: Czuba, Huk, & Cormack, 2011), and ultimately be evaluated for correspondence with single-neuron recordings.

Chapter 4: To CD or not to CD: Is there a 3D motion aftereffect based on changing disparities?[8]

The perception of motion through depth has received renewed attention lately, due to an accumulation of evidence that a binocular mechanism for such 3D motion uses motion signals per se as its input primitive (as opposed to the spatial position signals used by disparity-based stereopsis). Two articles on this topic, one by Czuba et al. (2011) and one by Sakano, Allison, and Howard (2012) were recently published in this journal. Because the two papers are so strikingly similar, we thought it would be worthwhile to compare and integrate the two sets of results.

Both papers used adaptation paradigms to establish the presence of a motion aftereffect (MAE) through depth (i.e. directly towards or away from the observers), and both papers used similar stimulus manipulations to isolate the contributions of the disparity-based cue (changing disparity, or "CD") from the velocity-based cue (interocular velocity difference, or "IOVD"). One difference in methodology was that Sakano et al. (2012) used the subjective duration of the MAE and the percentage of trials on which the MAE was reported as their dependent measures, whereas Czuba et al. (2011) used a 2AFC motion-nulling paradigm that yielded full psychometric functions from each observer for each condition. A second difference was that, in our original experiments, we adapted observers to large-field unidirectional 3D motion moving toward or away from the observer as shown on the left side of figure 4.1A. In contrast, Sakano et al. focused on a bi-directional adaptation stimulus consisting of oppositely moving stimulus regions above and below fixation.

Crucially, both papers found strong evidence for an IOVD mechanism that compares velocity signals between the two eyes to extract three-dimensional motion. This, along with a spate of recent papers by several other groups (Brooks, 2002b; e.g.

---

[8] This chapter is based on:
Czuba, T. B., Rokers, B., Huk, A. C., & Cormack, L. K. (2012). To CD or not to CD: Is there a 3D motion aftereffect based on changing disparities? Journal of Vision, 12(4). doi:10.1167/12.4.7

Fernandez & Farell, 2006; Nefs, OHare, & Harris, 2010; Shioiri et al., 2008; 2009) is an important convergence of evidence because the IOVD mechanism is a radically different kind of binocular computation; one that is distinct from the long-appreciated mechanism of stereopsis, and its confirmed existence has important theoretical and practical consequences for vision science.

However, one critical discrepancy between the two papers is that Sakano et al. (2012) found no evidence for a disparity-based (CD) contribution to the 3D MAE, whereas Czuba et al. (2011) did find that a 3D MAE could be produced by changing disparities, albeit one that was much smaller than that produced by IOVDs. To determine whether either of the two methodological differences mentioned above could explain our disparate findings, we conducted a follow-up experiment in which observers adapted to both towards and away motion simultaneously (Figure 4.1A, right side), in the spirit of the bi-directional geometry of Sakano et al. As in our previous experiment, we employed a 2AFC motion-nulling methodology to measure the magnitude of the resulting motion aftereffects. Adaptation stimuli consisted of random dot stereograms moving at 0.6°/second in opposite directions in the two eyes—generating percepts of 3D motion directly toward or away from the observer. The CD cue was isolated by dynamically updating dot positions on every display frame (60Hz) while maintaining steadily changing disparities consistent with a plane of dots moving through depth. The 'Full' Cue stimulus consisted of coherently-moving binocularly-paired dots, which inherently contained both CD and IOVD cues to 3D motion (for further details and stimulus examples see Methods section of Czuba et al., 2011). The only departure from our previously published methods was that test stimuli were randomly presented in either the upper or lower –but not both– hemifield on every trial. The logic is that the presence of an effect in this follow-up experiment would indicate that the 2AFC motion-nulling methodology is required to reveal the smaller CD MAE. Conversely, the absence of an effect would imply that adapting to a large uni-directional CD-isolating stimulus yields measurable CD-based aftereffects that were not present when adapting to different directions of motion in different parts of the visual field.

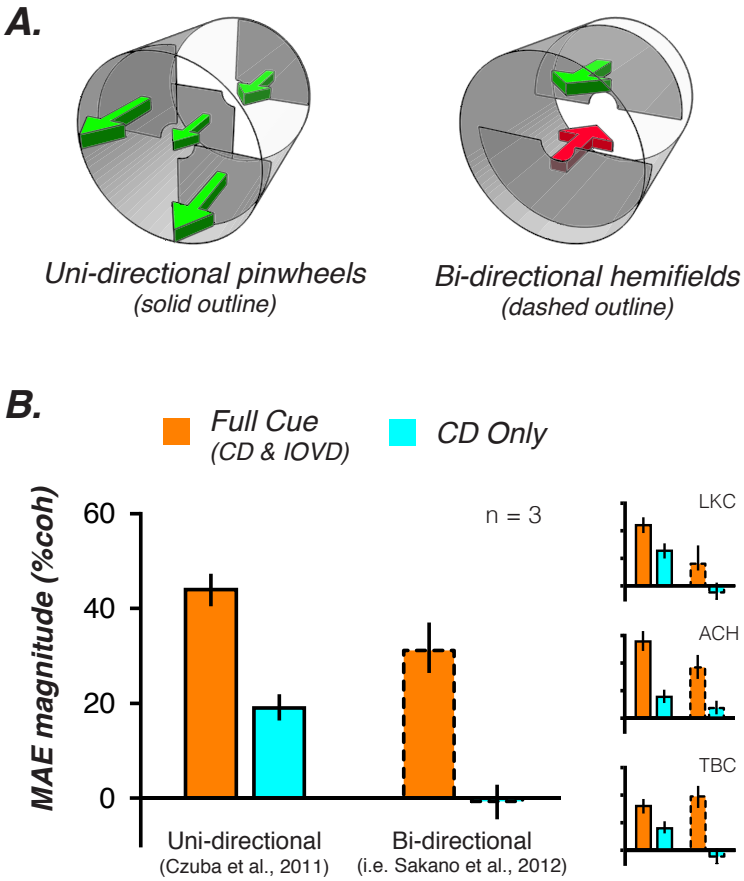

Figure 4.1 MAE following unidirectional or bidirectional adaptation

(A) Schematics of uni-directional adaptation stimuli used by Czuba et al. (2011) and bi-directional stimuli similar to Sakano et al. (2012). (B) Bar graphs of MAE magnitude, as estimated from the test motion coherence at which observers were equally likely to report seeing toward or away motion [the point of subjective equality]. Orange bars correspond to 'Full Cue' adaptation stimuli containing both binocular cues ('3D-planar' condition of Czuba et al.; 'RES' condition of Sakano et al.). Cyan bars correspond to CD-isolating adaptation stimuli ('CD' condition of Czuba et al.; 'DRES' condition of Sakano et al.). Error bars represent 95% confidence intervals on the bootstrapped distribution. Individual subject data are shown in the three smaller bar graphs to the right of the main figure.

The results of this experiment are shown in Figure 4.1B, which compares them to the corresponding data from our original 3D MAE paper. Bar graphs depict MAE magnitudes, as measured from the shift in point of subjective equality following adaptation. MAE magnitudes resulting from uni-directional adaptation toward or away from the observer are shown on the left (solid outlines; replotted from figure 13 of Czuba et al., 2011), while the corresponding MAEs from bi-directional adaptation (dashed outlines) are shown on the right side of the graph. Clearly, there remains a strong 3D MAE for Full Cue stimuli (orange bars) that contain both IOVD and CD cues, but the CD-isolated MAE (cyan bars) is now absent.

We note that Sakano et al. also did not find a CD MAE using uni-directional stimuli in a later experiment (Expt. 3, stimulus Case 2 & 3). However, the moving portion of those displays were very small, and were placed on a large, static, disparity pedestal (see figure 7 of Sakano et al., 2012). The constellation of these results points to a CD mechanism that requires both large and uni-directional 3D motion to elicit a motion aftereffect. The presence of a CD MAE, albeit under limited conditions, confirms that a stereomotion display with no coherent monocular motions can still generate directional signals in the visual system (Norcia & Tyler, 1984).

In conclusion, two independent groups of investigators used very similar psychophysical adaptation paradigms to reveal the presence of a mechanism that uses the difference in velocity between the two eyes to compute 3D motion. This IOVD mechanism is distinct from classical stereopsis in that it does not rely on fine positional differences between the eyes (and hence does not need to solve the classical "correspondence problem"). There was a disagreement, however, about whether the CD cue is able to produce a 3D MAE (and, presumably, contributes to 3D motion perception as a true directional signal). Comparison of our original and follow-up experiments suggests that 3D MAEs from isolated CD cues can be elicited with large uni-directional stimuli, but are much weaker (or even absent) for bi-directional and/or smaller adaptation stimuli. Whether this implies surprisingly large spatial summation in the sensory mechanisms that encode CD (Brooks & Stone, 2006b), capacity or resolution limits in

higher-level mechanisms, or other factors, is a current topic of work in our laboratory. Regardless of the etiology of the dependence of CD MAEs on spatial stimulus properties, the delicate nature of MAEs based on changing disparities (once thought to be the primary binocular cue to 3D motion) further emphasizes the renewed focus on the IOVD cue.

Chapter 5: Neural locus of 3D motion processing

INTRODUCTION

The significance of the velocity-based cue to 3D motion perception motivates the amendment of canonical models of motion processing to include the comparison of eye-specific motion signals. To better understand how this computation fits into known motion processing pathways we are currently exploring the physiological basis for distinct representation of 3D directions of motion using functional magnetic resonance imaging (fMRI) techniques. Previous work has implicated extrastriate areas in and around human MT+ (Rokers et al., 2009; Likova & Tyler, 2007; respectively) as being involved in 3D motion processing, yet it is unclear whether this is the result of a unified '3D motion' mechanism or distinct disparity- and velocity-based mechanisms. Using event-related fMRI adaptation protocols (Huk et al., 2001; Larsson et al., 2006) we are able to compare the neural activation resulting from isolated 3D motion cues to better understand how the visual system might combine multiple sources of 3D motion information.

Previous research has shown that 2D motion adaptation causes a directionally selective decrease in fMRI BOLD signal in human MT+, a key structure involved in visual motion processing (e.g. Maunsell & Van Essen, 1983; Mikami, Newsome, & Wurtz, 1986; Rust et al., 2006), as well as smaller yet significant effects across a range of intermediate and early visual areas leading up to MT+ (Huk et al., 2001; H. A. Lee & Lee, 2012). More recently, Rokers et al. (2009) measured fMRI activation following adaptation to 3D motion stimuli containing both velocity- and disparity-based motion cues. Direction-dependent adaptation effects were seen most strongly in MT+, and to a lesser degree in intermediate and early visual areas V2, V3, V3A, and LO, but surprisingly not in V1. Relative to 2D motion adaptation, 3D adaptation effects were significantly stronger in LO and MT+. Area LO is a visually responsive region of human lateral occipital cortex located between dorsal V3 and MT+. Distinct functional properties of LO are not well understood, but it has been implicated in the processing of

motion/object boundaries and structure-from-motion stimuli (Larsson & Heeger, 2006; Read, Phillipson, Serrano-Pedraza, Milner, & Parker, 2010; Tyler et al., 2005; Zhuang, Peltier, He, LaConte, & Hu, 2008).

Building upon psychophysical evidence for distinct representation of 3D directions of motion, this chapter examines physiological evidence for binocular 3D motion processing via functional neuroimaging. Following adaptation to unidirectional IOVD or CD cue isolating 3D motion stimuli, directionally-selective adaptation effects were measured by comparing the fMRI BOLD response during presentation of brief probe stimuli moving in either the same or opposite direction as adaptation. Regions of cortex that are selective for 3D motion should exhibit a decreased BOLD response to probe stimuli presented in the *same* direction as adaptation, relative to the *opposite* direction (Huk et al., 2001; Mather, 1980; Rokers et al., 2009). Experimental stimuli were designed to match those used in psychophysical adaptation experiments (Chapter 3) as closely as possible (e.g. cue isolation, timecourse of top-up adaptation & test). Neuroimaging was conducted using an event-related adaptation protocol developed by Larsson et al. (2006), which utilizes 'blank' trial activation subtraction to isolate the responses to brief probe stimulus presentations (see Methods & Figure 5.2).

This neuroimaging component addresses two important aspects of binocular motion mechanisms that cannot be gleaned from the previous psychophysical results. First, whether distinct cortical circuits are involved in CD and IOVD computations. Representations of retinal disparity and (2D) motion are found extensively throughout the dorsal visual stream (Neri, Bridge, & Heeger, 2004; Orban, Kennedy, & Bullier, 1986; G. F. Poggio, Gonzalez, & Krause, 1988; Priebe, Lisberger, & Movshon, 2006). However, 3D motion adaptation using stimuli that contained concurrent disparity- and velocity-based 3D motion cues revealed 3D direction selectivity in a limited subset of extrastriate visual areas: MT+, LO, & V3A (Rokers et al., 2009). Given this distribution of 3D direction selectivity, isolated cue adaptation may reveal differential recruitment of cortical areas depending on the disparity- or velocity-based computations involved.

Regardless of whether disparity- and velocity-based cues are computed by separate mechanisms, the question of whether a unified representation of 3D motion remains. Further examination of cross-cue adaptation effects will help to determine whether disparity- and velocity-based adaptation effects reflect a multiple, distinct, mechanisms for binocular motion processing, or a unified representation of 3D directions of motion.

METHODS

Subjects

fMRI data were collected from three subjects (all males, aged 29–*49*), all with good stereopsis, and normal or corrected-to-normal vision. Subjects were the same three experienced psychophysical observers who participated in the 3D motion adaptation experiments described in Chapter 3. Experiments were undertaken with the written consent of each subject and procedures were approved by the UT–Austin Institutional Review Board and the Imaging Research Center safety guidelines. Each subject completed 12 functional-MRI scanning sessions.

Functional imaging & processing

Magnetic resonance imaging was acquired at The University of Texas at Austin Imaging Research Center using a GE Signa HD 3T scanner with an 8-channel phased array head-coil. Subjects were scanned using a two-shot spiral BOLD fMRI sequence (3.2 mm isometric voxels, 1.5 s volume acquisition, TR = 750 ms, TE = 30 ms, flip-angle of 56°), with 14 slices oriented parallel to the calcarine sulcus. Each scanning session consisted of 10 scans, lasting 270 seconds (180 scanner frames) each. The first and last 15–30 s of each scan (two trials for event-related scans, one block for localizer scans; respectively) were discarded to allow for transient scanner start-up effects and hemodynamic saturation.

Each scanning session began with a T1-weighted anatomical localizer (26 sagittal slices) which was used to align the functional slice prescription parallel to the calcarine sulcus. An additional T1-weighted anatomical scan was performed to acquire anatomical in-plane images for each of the 14 slices in the functional imaging prescription. These in-plane anatomies were later used for co-registration of scanning sessions and cortical regions of interest (ROI). Foam padding was used to stabilize the subject's head throughout each session. Functional data were corrected for movement both within and between scans (Nestares & Heeger, 2000). Each fMRI time-series was high-pass filtered and demeaned to compensate for slow signal drift and to convert arbitrary image intensity units into percent BOLD signal change. Individual voxel time-series within each ROI were then averaged together and restricted to grey matter that was responsive to functional localizers conducted during the first and last scan of each session. All fMRI signal processing and analyses were performed using standardized methods implemented in Matlab (The MathWorks, Natick, MA) by the Stanford mrVista software package (http://vistalab.stanford.edu/mrvista).

In-plane anatomies were co-registered to grey-matter segmented high-resolution (1 mm isometric) whole-brain anatomical volumes. Mapping of visual areas V1, V2, V3, V3A, V4, LO, MT, & MST were performed in separate sessions for each subject following standard retinotopic mapping procedures (Dougherty et al., 2003 and see Rokers et al. 2009). For two of the three subjects additional functional reference sessions were conducted to distinguish MT & MST sub-regions within MT+ using methods established by Huk et al. (2002). In the third subject, purely retinotopic methods were used to delineate MST as an anterior region of MT+ containing a distinct retinotopic map with activation corresponding to the ipsilateral visual field.

Visual stimuli

The first and last scans of each session were functional MT+ reference scans that consisted of stimulus block alternation (15 s on, 15 s off) of moving versus static dot fields. Building on classic two-dimensional motion MT+ reference stimuli (Tootell et al.,

1995; J. D. Watson et al., 1993; Zeki et al., 1991), we used a dynamic 3D motion stimulus that consisted of a coherently moving dot field that pseudo-randomly swirled through a three-dimensional volume equal to that of the experimental stimulus. Reference scans were used to further constrain ROIs to voxels that were significantly activated by excluding all voxels with a covariance threshold of < 0.3, an estimate of signal-to-noise based on the best fit sinusoid at the stimulus alternation frequency.

The remaining eight scans in each session employed an event-related adaptation protocol to measure the effect of unidirectional 3D motion adaptation on the BOLD response to test stimuli presented in the same or opposite direction as adaptation. A single adaptation direction (toward or away) and motion cue type (Full, IOVD, or CD) was pseudo-randomly selected, and maintained throughout all adaptation scans in a session (i.e., the 2nd–8th scan of each session). BOLD fMRI responses were measured while observers viewed a sequence of stereoscopically presented moving dot adaptation and test stimuli that were largely identical to those described in chapter 3. Prior to the first adaptation scan, a steady state of 3D motion adaptation was established by presenting 60-seconds of 3D motion adaptation (i.e. pre-scan adaptation, Figure 5.1). Within each scan, a series of trials were presented, where each trial consisted of top-up adaptation (4 s), followed by a brief probe stimulus presentation (1 s; 1.25 s inter-stimulus interval), moving in the Same or Opposite direction as adaptation (from here on referred to collectively as a 'trial'). Each trial lasted a total of 7.5 seconds, or 5 fMRI acquisition frames, and 36 trials were presented in each session (270 s total).

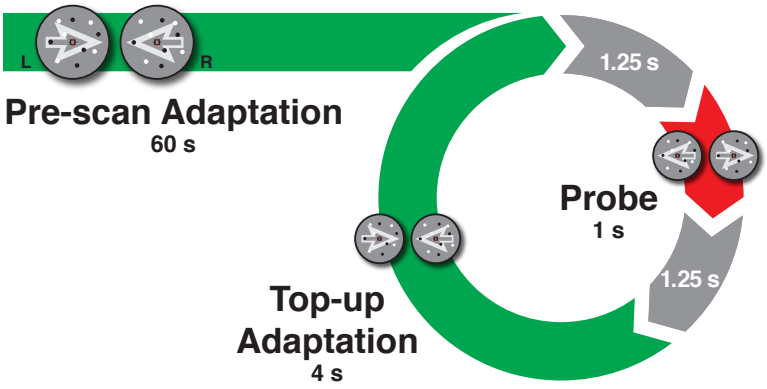

Figure 5.1 Schematic of event-related trial sequence

Time course of adaptation and probe presentation is nearly identical to the psychophysical 3D adaptation paradigm used in Chapter 3. A pre-scan adaptation (60 s) was only presented prior to the first scan. Each trial consisted of 4-second top-up adaptation followed by a 1-second probe stimulus presentation in either the same or opposite direction as adaptation (in this example, the opposite direction), with 1.25-second inter-trial intervals.

By design, very few adjustments were necessary between psychophysical and neuroimaging applications of our adaptation protocol. (1) In order maximize observable BOLD response, test stimuli consisted of fully coherent (i.e. 100% signal dot) stimuli moving in either the same or opposite motion as adaptation stimuli, rather than the broad sampling of coherences used to measure psychometric functions. (2) Inter-stimulus intervals were extended from 0.25 s to 1.25 s to coincide with the timecourse of fMRI volume acquisition. (3) On 22% of trials (8 of 36 trials per scan) test stimuli were omitted and only a mean grey background was presented. Following adaptation protocol developed by Larsson et al. (2006; also used in Rokers et al., 2009), subtraction of the BOLD time series during these 'blank' trials allowed us to visually isolate the BOLD response to the brief probe stimulus presentation.

Because each 270-second adaptation scan was presented in rapid succession (~4 s between scans), the 60-second pre-scan adaptation stimulus was only presented after the first MT reference scan; before the first adaptation scan. Steady-state motion adaptation was confirmed by comparing direction selectivity indices based on the first- and last-four scans of each session.

The final distinction between psychophysical and neuroimaging experiments lies in the necessity to control for confounding factors of subjects' attentional state during functional imaging. Huk et al. (2001) showed that the direction-dependent decreases in BOLD signal caused by motion adaptation are easily washed out, or even inverted, due to increased attention to the illusory motion percepts caused by motion adaptation. Properly controlling for attention reveals a directionally-selective *decrease* in BOLD response following adaptation, supporting the notion that the motion aftereffect is the result of an imbalance in directionally-selective neural populations (Mather, 1980). Throughout adaptation scans subjects performed an attentionally demanding two-interval forced choice (2-IFC) colored dot discrimination task in which they were asked to discriminate the interval with an unequal ratio of red and green dots. Each interval contained 30 non-overlapping red and green dots (0.2° dia.) presented in a 0.5° annulus eccentric to the experimental stimulus. The duration of each presentation interval was randomly selected

from a truncated exponential distribution (raised-cosine dot-wise onset/offset; t_μ = 2.0 s, 0.5 ≤ t ≤ 4.0 s). A yellow-dot presented at fixation cued the subject to report the perceived unequal colored-dot interval using a two-button response box, with a 2 s inter-trial interval. After each response, feedback was provided by displaying a green (correct) or red (incorrect) dot at fixation. Task difficulty was adjusted using an adaptive staircase procedure {QUEST, Watson:1983wd} to obtain approximately 82% correct performance. The motivation for variable, gradual onset times was three-fold: (1) to minimize abrupt onset transients, (2) to maintain subjects' attention within & between presentation intervals, and (3) to maintain temporal independence between the attentional control task and the experimental stimulus presentation.

Display apparatus

Pilot experiments revealed the otherwise standard fMRI projection display equipment (which provided a maximum stimulus eccentricity of just over 1°) to be problematic for 3D motion adaptation, and a larger stimulus area was necessary to allow direct comparison of effects across psychophysical and neuroimaging experiments. We therefore developed custom display similar to that used in the Speed & Eccentricity experiments (Chapter 2). Stimuli were presented on a large (42") consumer brand LCD television display that was modified to remove any extraneous ferromagnetic components; replacing virtually all of the external enclosure with a custom built plexiglass frame which was then enclosed in a windowed Faraday box. In combination with a custom MR safe mirror stereoscope, this hardware setup enabled unparalleled accuracy in 3D stimulus presentation during our fMRI experiments.

fMRI data analysis

The fMRI time series for each trial was defined as the 22.5 s period (3 adapt-probe presentation sequences) following the onset of the top-up adaptation stimulus. Each trial time series spanned three adaptation–probe presentations in order to accommodate the slow hemodynamic delay of the BOLD response. To maintain an equal number of

time points for each trial, the last two trials of each scan were discarded. Trial time series for each ROI were then averaged across voxels and within trials based on whether the probe stimulus moved in the same or opposite direction as adaptation. The response to the probe stimulus was then isolated by subtracting the mean time series of the 'blank' trials from the 'same' and 'opposite' direction time series (Figure 5.2).

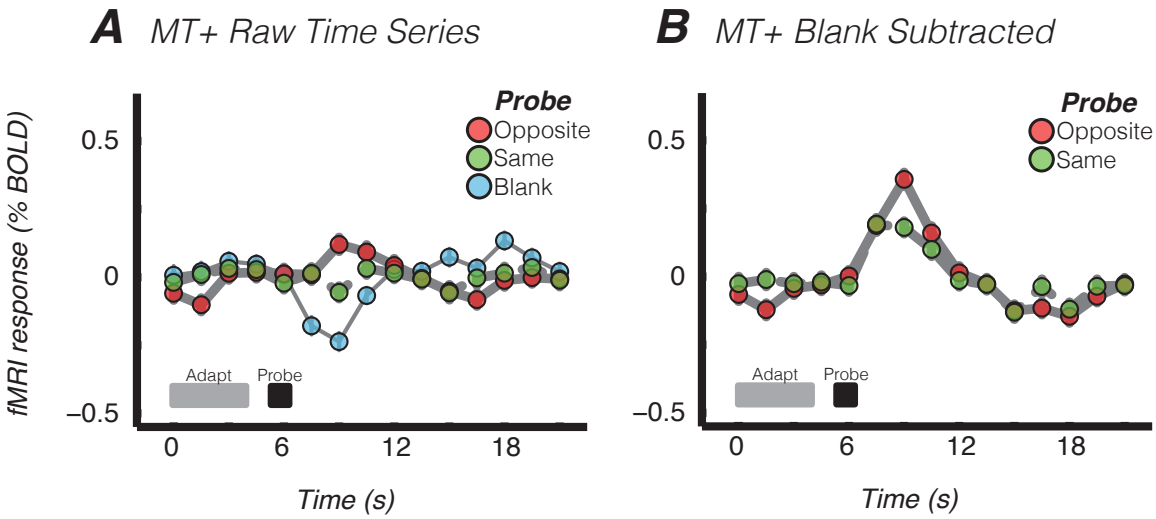

Figure 5.2 Example functional time series & blank subtraction

MT+ response following adaptation to Combined Cue 3D motion (i.e. both IOVD & CD). (A) Time series of mean fMRI BOLD response (y-axis) aligned to onset of the top-up adapter (grey bar). (B) The response to probe stimulus presentation (black bar) can be isolated by subtracting Blank time series responses (blue) from trials in which the probe moved in the Opposite (red) or Same (green) direction as adaptation. Time series represent mean response averaged across 3 subjects.

Within each scanning session, univariate response amplitudes for each trial were computed by projecting the trial time series onto the overall mean time series (regardless of probe direction) and normalizing by the amplitude of the mean response:

$$A_i = \frac{R_i \cdot \overline{R}}{||\overline{R}||} \qquad \text{(eq. 4.1)}$$

Where R_i is the time series vector of the i^{th} trial, $\overline{R}$ is the mean time series vector, and A_i is the scalar response amplitude in units of % BOLD signal change. Individual trial amplitudes were then averaged for each probe direction and used to compute a direction selectivity index:

$$I_{Dir.Sel.} = \frac{\overline{A}_{opp} - \overline{A}_{same}}{|\overline{A}_{opp}| + |\overline{A}_{same}|} \qquad \text{(eq. 4.2)}$$

Where $\overline{A}_{opp}$ and $\overline{A}_{same}$ are the mean projected amplitudes of probe trials moving in the opposite and same direction as adaptation, respectively, and $I_{Dir.Sel.}$ is the direction selectivity index for a given ROI. Direction selectivity indices ranged from –1 to 1, where a positive direction selectivity index corresponds to decreased BOLD response following adaptation.

PRELIMINARY RESULTS

Building on previous evidence that neurons in area MT+ are involved in not only 2D but also 3D motion processing, we measured fMRI responses in the visual cortex while subjects observed a series of 3D motion adaptation and probe stimuli designed to isolate velocity- or disparity-based cues to 3D motion. Although psychophysical experiments have shown evidence for IOVD and—under the right conditions—CD based representation of 3D directions of motion, questions remain about how binocular mechanisms with such distinct functional characteristics could be integrated into a unified representation of 3D motion.

As a baseline, we first measured fMRI adaptation to 3D motion stimuli that contain CD and IOVD cues to 3D motion. Figure 5.3 plots the mean time course of fMRI

responses from 3 observers for a range of visual areas leading through V1, up the dorsal stream, to area MT & MST. Clear directionally selective 3D motion responses can be seen in the time course of fMRI responses in areas MT & MST; evidenced by a greater response to opposite-direction probes than those that moved in the same-direction as adaptation. Despite a wealth of disparity and 2D motion sensitivity in primary visual cortex, very little direction selectivity can be seen outside of MT and MST.

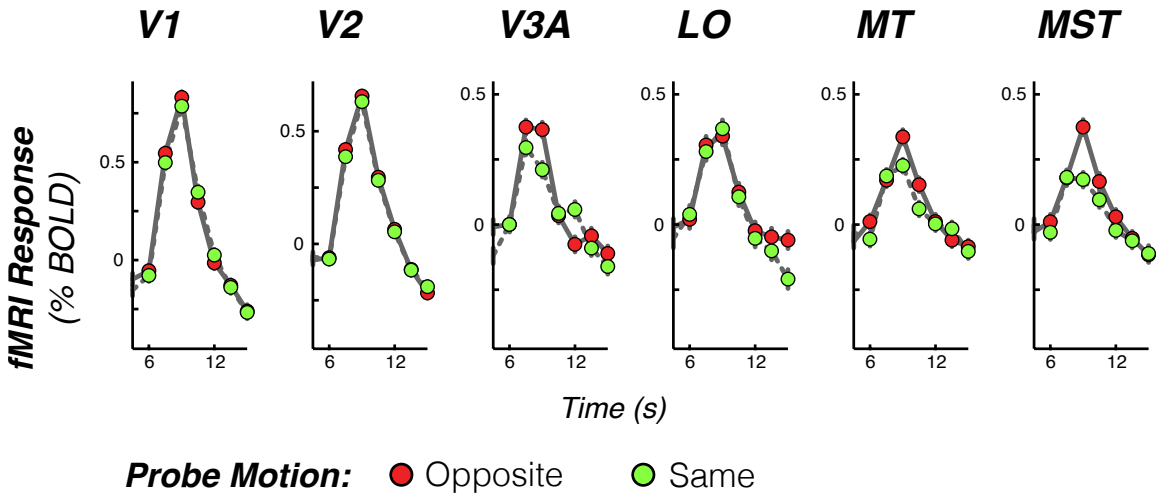

Figure 5.3 Combined cue adaptation across visual ROIs

Time series of fMRI responses following adaptation to 3D motion stimuli containing *both* binocular cues. As in Figure 5.2b, red and green time points correspond to probe trials moving in the opposite or same direction as adaptation. Peak responses to opposite-direction trials are significantly higher that same-direction trials in MT and MST, but not in early early visual areas.

To more quantitatively address 3D direction selectivity across visual cortex, we computed direction selectivity indices for each ROI based on the projected amplitudes of the probe responses to each motion condition (Figure 5.4). A complimentary series of fMRI measurements to examine how adaptation to isolated IOVD (middle row) and CD (bottom row) cues affect the pattern of direction selectivity observed from combined cue 3D motion adaptation are shown in the middle and bottom row of figure 5.4. While these are only preliminary results, the pattern of direction selectivity from IOVD and combined cue adaptation only emerging within MT+ not only supports previous findings, but also provides encouraging evidence that neurons in MT+ are not only involved in 2D motion computations, but also appear to be the neural locus of the interocular velocity difference computation. Meanwhile, adaptation to CD isolating stimuli produces considerably weaker direction selectivity indices across the visual cortex. Although it is tempting to delve into the finer details of the relative 3D direction selectivity within subunits of MT+, curious patterns of selectivity in intermediate areas (e.g. V3A & LO) that are relatively less understood than MT or V1, the preliminary nature of these results do not yet have the statistical power to make any concrete inferences about underlying interactions of binocular 3D motion mechanisms. However, it seems clear that much of the directional selectivity observed from combined cue adaptation is being driven by the IOVD mechanism.

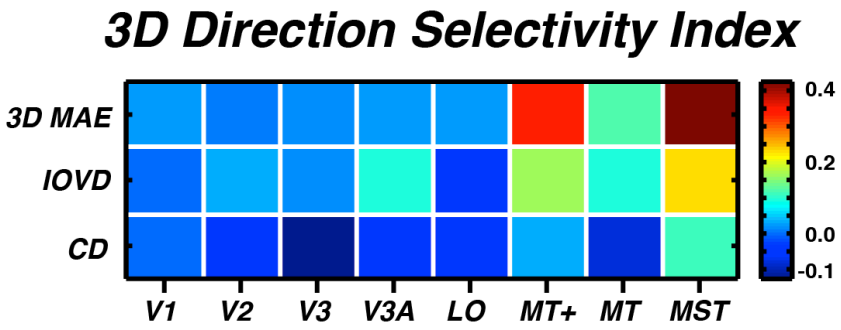

Figure 5.4 3D direction selectivity index

Direction selectivity indices from adaptation to isolated binocular motion cues, computed at a range of functional ROIs along the dorsal visual pathway.

Beyond direct comparison of isolated-cue adaptation effects across visual areas, our fMRI adaptation protocol offers the chance to explore cross-cue adaptation effects. Cross-adaptation has been used in a number of visual modalities to determine whether certain stimulus characteristics share a common underlying mechanism or are separately processed by the visual system (e.g. Anstis, Giaschi, & Cogan, 1985; Durgin & Hammer, 2001; Georgeson & Schofield, 2002; Pianta & Gillam, 2003). Likewise, cross-adaptation of binocular 3D motion cues will provide important insight on whether the disparity- and velocity-based cues are processed separately, or by a unified 3D motion mechanism.

Chapter 6: General Discussion

The perception of 3D motion is undoubtedly highly influenced by binocular motion information, yet the specific mechanisms involved are still relatively poorly understood. The tension between IOVD and CD cues lies in the fact that under normal viewing conditions the two cues coexist and encode the same 3D direction information, yet each mechanism has distinct implications for how and where 3D motion is processed. In the simplest of models, their differences lie in the relative order of temporal and binocular integration. The disparity-based computation implies an early-stage binocular combination followed by temporal integration of disparity signals, while the velocity-based cue posits early temporal integration by monocular velocity signals, followed by binocular combination through interocular comparison of velocity signals in corresponding regions of the two eyes.

The preceding series of experiments paint a surprising picture of 3D motion processing that relies primarily on a mechanism based on comparisons of interocular velocity signals, with little dependency on concurrent changes in binocular disparity. The following section discusses these findings in relation to previous research, addresses some defining features and limitations of the CD and IOVD cues, and summarizes some ongoing aspects of our research and unresolved questions in binocular 3D motion processing.

RELATION TO PAST WORK

By deconstructing 3D motion cues into their fundamental components, important insights can be gained about how specific binocular motion cues are be more or less useful under real world conditions. The oft cited example of a baseball or cricket player's ability to hit a rapidly approaching ball, without being hit himself, is perhaps an overly constrained scenario when it comes to binocular motion computations (Cynader & Regan, 1978; Regan & Gray, 2000). Although the ability to perceive/predict the trajectory of an object with such amazing accuracy and swiftness is an impressive task, a great deal

of prior knowledge about the spatial and temporal aspects of the object's motion is already known. An observer that can orient their gaze to the proper location in space and time prior to the event in question might have a better chance of detecting small changes in the object's disparity (Land & McLeod, 2000). However, a better example of 3D motion detection might be that of a basketball player reacting to an unexpected pass. The player must detect an approaching object amongst a highly noisy motion environment with relatively little prior knowledge of the spatial and temporal origins of the pass. In such cases it is necessary to detect relevant 3D motion originating from an unknown location in the environment that is likely outside of the limited range of binocular fusion (Blakemore, 1970; Schor & Badcock, 1985; Westheimer & Truong, 1988). Poor disparity sensitivity at locations that are far from fixation, in either eccentricity or in depth, make the CD mechanism an ill suited candidate for detection of 3D motion throughout the visual field. By side-stepping the binocular correspondence problem (Burt & Julesz, 1980; G. F. Poggio & Poggio, 1984), the interocular velocity difference cue provides a spatially robust mechanism for detecting 3D motion, even in the absence of strict binocular correspondence. Shioiri et al. (2000) found that observers were able to accurately discriminate the direction of 3D motion when presented binocularly unmatched arrays of dots (i.e. spatially uncorrelated arrays of dots presented to each eye) moving in opposite horizontal directions in the two eyes. Such a scenario could realistically arise from 3D motion of a large object located beyond the range of binocular fusion, and would be an important class of stimuli for the visual system to detect.

Due to the longstanding inability to isolate the IOVD cue, many studies were limited to inferences based on sensitivity differences between CD-isolating and combined cue (CD & IOVD) stimuli (e.g. Portfors-Yeomans & Regan, 1996). More recently, a number of techniques have been developed to effectively isolate interocular velocity differences from concurrent changing disparity signals (Fernandez & Farell, 2005; Rokers, Cormack, & Huk, 2008b; Shioiri et al., 2000). These methods have allowed a more balanced assessment of the relative utility of IOVD and CD cues and have led to a resurgence of interest in IOVD contributions to 3D motion processing.

Although many recent findings suggest a more significant role for the IOVD cue than previously thought, important distinctions exist between the psychophysical stimuli and tasks have been used in efforts to disentangle the binocular mechanisms involved. For instance, the aspects of 3D motion an experimenter considers important to performance (e.g. velocity dependence, spatial resolution, direction discrimination), and their resulting metric of sensitivity can have significant influence on the appearance of relative binocular cue contributions. Experiments that have reported strong contributions from the disparity-based cues often consist of relatively small stimuli presented very near the fovea, and utilize tasks designed to measure minimum signal detection or displacement thresholds (Cumming & Parker, 1994; Gray & Regan, 1996; Portfors-Yeomans & Regan, 1997). On the other hand, studies reporting significant contributions from velocity-based cues more often arise from measurements of direction and speed discrimination sensitivity using stimuli that are presented outside the fovea and have a generally broader spatial extent (Brooks, 2002b; Brooks & Stone, 2006a; J. M. Harris & Watamaniuk, 1995; Rokers et al., 2009). Rather than passing off these differences as experimental artifact, it is possible that the visual system is capable of utilizing different sources of binocular information depending on the situation at hand (e.g. M. L. Morgan, Deangelis, & Angelaki, 2008).

Following this logic, we designed our experiments with the goal of determining the conditions in which individual binocular cues can best account for 3D motion sensitivity when both cues are available; as opposed to reverse engineering combined cue performance based on the limits of isolated motion cue stimuli. We believe this framework provides the best likelihood of obtaining an unbiased estimate of how binocular motion cues are utilized under normal viewing conditions and is an approach that has been increasingly embraced as methods of experimentally isolating binocular motion cues have improved (Brooks & Stone, 2004; Shioiri et al., 2008).

Additionally, all of our experiments were designed with significant influence from 2D motion literature and manipulations of motion coherence (Blake & Hiris, 1993; Newsome & Paré, 1988; Watamaniuk et al., 1989), more so than positional stereopsis and

manipulations of stereo correspondence (Cormack et al., 1991; Palmisano, Allison, & Howard, 2006). (Cumming & Parker, 1994; Gray & Regan, 1996)

Electrophysiology

In traditional electrophysiological recordings, direction selectivity is almost always measured by an iterative process of adjusting stimulus position, size, and speed to determine initial stimulus parameters that produce a qualitatively maximal response, followed by a more principled measurement of neural responses to a range of directions of motion in order to build up a representation of the neuron's directional tuning (Britten, Shadlen, Newsome, & Movshon, 1992; Perge, Borghuis, Bours, Lankheet, & van Wezel, 2005). Likewise, many electrophysiological attempts to study binocular 3D motion mechanisms begin by characterizing peak monocular velocity tuning before examining interactions in relative orientation tuning in the two eyes (Cynader & Regan, 1978; e.g. Zeki, 1974). However, a direct extension of traditional approaches for determining 2D direction selectivity may not be the correct approach for detecting neural tuning for 3D motion. One important distinction between 2D and 3D motion is that relatively slow monocular speeds give rise to brisk 3D motion speeds. Because the velocity tuning to frontoparallel 2D motion in MT tends to be centered around 8-16°/s (J. Liu & Newsome, 2003), presenting opposite monocular velocities at peak monocular response speeds would correspond to unrealistically fast 3D motions. For example, from a 3 m viewing distance, equal and opposite monocular velocities of 12°/s are equivalent to a jet flying at Mach 1 directly toward the observer. Under such circumstances, either velocity or disparity cues would seem to be woefully ineffective for directing any actionable response. On the other hand, a neuron sensitive to 3D directions of motion could exhibit distinct orientation tuning functions for high and low monocular velocities, as well as binocular interactions that are only apparent during stereoscopic stimulus presentation (Akase et al., 1998; G. F. Poggio & Talbot, 1981). It is therefore crucial that efforts to examine the neural representation of binocular 3D motion cues employ stimuli that contain monocular velocities that correspond to plausible 3D motion speeds.

In one of the first electrophysiological endeavors specifically aimed at studying the neural representation of 3D motion, Zeki (1974) measured the direction tuning of binocularly driven MT neurons to determine whether distinct monocular tuning existed for motion presented in the two eyes. Unfortunately, the primitive nature of stimulus presentation (hand-held projectors were used to manually sweep a bar of light across the projection screen) it was not possible to binocularly simulate 3D motion (i.e. directional tuning was determined in separate monocular presentations) and optimal stimulus speeds were not specifically reported for each cell. Never the less, a very small percentage of neurons sampled exhibited tuned responses to opposite directions of motion in the two eyes. Notably, the few cells that did exhibit opposite direction selectivity between the two eyes seemed to be clustered in small patches of MT. Neurons in MT have been shown to have a high degree of topographical organization in their sensitivity to direction and disparity (Deangelis & Newsome, 1999), but not speed (J. Liu & Newsome, 2003).

In another study, Poggio & Talbot (1981) provide the first extensive electrophysiological examination of selectivity for motion through depth using binocular stimulation and (primarily) speeds that correspond to plausible real world 3D motion speeds. Although their recordings were located in primary visual cortex (areas A17 & A18), they found a similarly sparse number of neurons with inter-ocularly opposite direction tuning. Notably, the use of binocular stimulation in Poggio & Talbot's treatment provided the ability to determine disparity sensitivity of putative 3D motion neurons, which were found to be largely insensitive to positional disparity.

Constraints on putative CD mechanisms

Despite a wealth of psychophysical evidence that 3D motion can be perceived based on changing disparity information alone, there is scant evidence for explicit neural representation of the temporal derivative of disparity. Selectivity for static disparity is a common feature throughout the visual cortex, leaving little question of whether inputs necessary for a CD mechanism are available (Neri et al., 2004; but see Ohzawa, DeAngelis, & Freeman, 1996). However, the prevalence of static disparity tuning has

actually hindered efforts to produce physiological evidence for a CD mechanism. Maunsell & Van Essen (1983) showed that previous electrophysiological findings that seemed to exhibit 3D motion selectivity could be better explained by a disparity dependent increase in response as 3D motion stimuli moved across a gradient of static disparity sensitivity (see Figure 8 of Maunsell & Van Essen, 1983). Therefore, neurons truly selective for 3D motion by means of a changing disparity mechanism should *not* be independently selective for static disparities.

Peng & Shi (2010) recently devised a physiologically plausible model for how a CD mechanism might be accomplished. By grafting static disparity energy computations (Ohzawa, DeAngelis, & Freeman, 1990) onto the inputs of the canonical 2D motion energy model (Adelson & Bergen, 1985; A. Watson & Ahumada, 1985) they present a Changing Disparity Energy Model (CDem) of 3D motion processing. Applying the CDem, dots moving toward or away from the observer at various speeds produces a bandpass 3D motion energy profile with peak sensitivity at 1 degree of disparity change per second (i.e. monocular velocity of 0.5°/s in our monocular motion frame of reference). Although this model provides an interesting first conceptualization of how a CD mechanism could be computed, in its current form, the CDem model does not reflect the low-pass speed/temporal frequency tuning for isolated CD cue stimuli or the higher peak sensitivity of combined cue sensitivity seen in psychophysical results (Czuba et al., 2010; Shioiri & Matsumiya, 2009).

In the process of addressing discrepancies between our 3D MAE results with those of Sakano et al. (2012), we were intrigued to find that the CD MAE only occurred with very large unidirectional adaptation stimuli. This supports previous findings that the CD cue is computed by a mechanism with much larger receptive fields than the IOVD cue (Brooks & Stone, 2006b). However, the lack of direction selective adaptation from neuroimaging of isolated CD adaptation motivated additional experiments examining the effect of attention on CD adaptation. Frontoparallel motion aftereffects induced with either luminance defined (M. J. Morgan, 2012) or disparity defined (Patterson et al., 2005) gratings have been shown to be robust to manipulations of attention. To test for an

interaction between attention and motion adaptation, we reexamined 3D motion aftereffects to stimuli identical to previous psychophysical experiments, while subjects simultaneously performed the attentional control task used in neuroimaging experiments. Although subjects reported the experience as being "torturous", preliminary results yielded measurable aftereffects in both IOVD and Combined Cue adaptation stimuli, but the CD MAE was totally absent when attention was diverted away from the adapting stimulus. Thus, the constellation of results from CD adaptation experiments suggest that *if* the temporal derivative of disparity is a quantity that is computed by the brain, it requires a surprising degree of overt top-down influence. It seems increasingly likely that 3D motion percepts arising from isolated CD cues are better described by a top-down inference of 3D motion, rather than an explicit neural computation of the temporal derivative of disparity.

Constraints on putative IOVD mechanisms

The research described herein builds on a growing body of evidence that comparisons of interocular velocities make up a significant contribution to binocular 3D motion processing. The importance of such a mechanism motivates the expansion current models of how motion signals from the two eyes are combined to form a unified visual representation of the three-dimensional world around us. It is typically assumed that separate monocular signals are combined into a single cyclopean representation at very early stages of visual processing. The presence of disparity selectivity as early as V1 (G. F. Poggio, Motter, Squatrito, & Trotter, 1985) and interocular transfer of monocular motion aftereffects (Blakemore & Campbell, 1969) indicate that the monocular visual pathways are combined to a single cyclopean form during relatively early stages of visual processing. Likewise, current models of binocular processing posit that by the time signals reach MT, most neurons are at least partially driven by signals from both eyes (Carney & Shadlen, 1992; Hess, Hutchinson, Ledgeway, & Mansouri, 2007; Maunsell & Van Essen, 1983). However, recent research suggests that the IOVD mechanism acts upon relatively late stages of motion processing.

A late stage IOVD computation is supported by psychophysical evidence that the visual system is capable of performing IOVD computations on monocular pattern motion signals. In the interest of examining the nature of motion signals that contribute to interocular velocity comparisons, we conducted a series of experiments to determine whether pattern motion signals contribute to binocular 3D motion processing, and if so, what computations underly the integration of component motion signals. Briefly, we presented drifting 120° plaid patterns (each component ±60 from pattern motion direction) that moved in opposite angular directions in the two eyes. By manipulating the angular direction of motion (maintaining 180° difference between eyes), we found that the magnitude and direction of 3D motion percepts varied as a function of the horizontal component of pattern motion (Figure 6.1); as opposed to that of individual component motions (see Appendix A for further details & conditions).

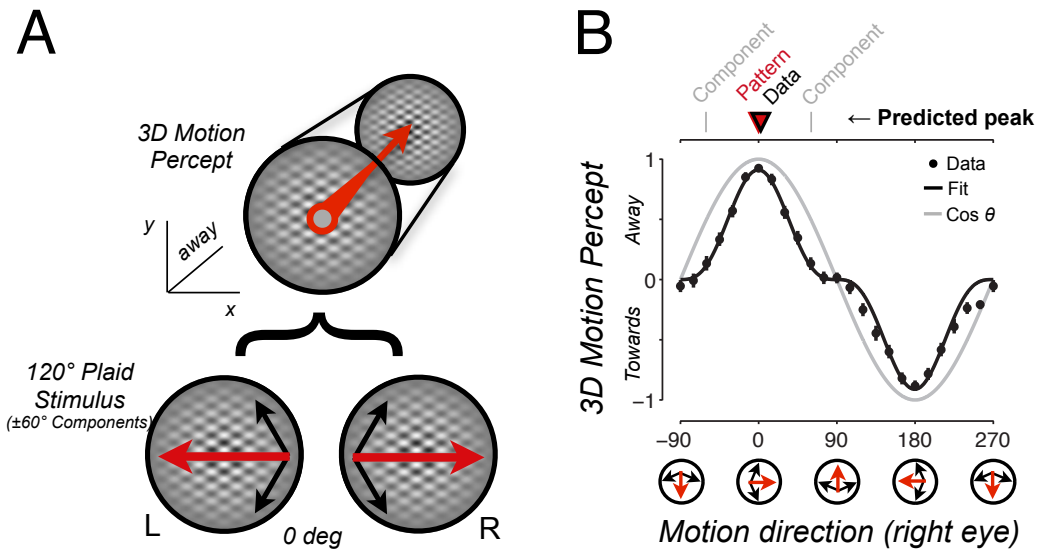

Figure 6.1 Binocular plaid (type I) schematic & results

(A) A schematic of binocularly presented 120° plaid stimuli (Type I). In each eye, two superimposed drifting gratings (black arrows, sf = 2 cyc/°, tf = 2 Hz) are presented with a relative orientation of 120°. Plaid patterns in each eye are oriented such that 180° angular difference between pattern motion directions (red arrows) is maintained. The arrangement pictured corresponds to the '0 degree' condition, corresponding to motion away from the observer. (B) Observers' reported perceived direction of motion (y-axis) as a function of the plaid motion direction in the right eye (x-axis; left eye always 180° opposite). Resulting 3D motion percepts are well fit by a cosine function (grey line) corresponding to the magnitude of horizontal pattern motion (i.e. strongest when pattern motions are in opposite horizontal directions). Adapted from Rokers, Czuba, Cormack, & Huk (2011); see Appendix A for original & complete details.

To determine whether resulting 3D percepts were derived from IOVD cues based on pattern motions or CD cues from plaid components (Delicato & Qian, 2005), we developed a novel binocular plaid stimulus that consisted of spatially-separated binocularly-unpaired gabor components; deemed 'pseudoplaids' (Figure 6.2; Clark & bradley, 2008). In total, the stimulus configuration provided a purely isolated IOVD signal that could only be computed based on monocular pattern motion signals that were integrated on a spatial scale greater than a single V1 receptive field (Appendix A for details). Indeed, subjects responses corresponded to 3D motion percepts that were in phase with the horizontal component of monocular pattern motion (Figure 6.3).

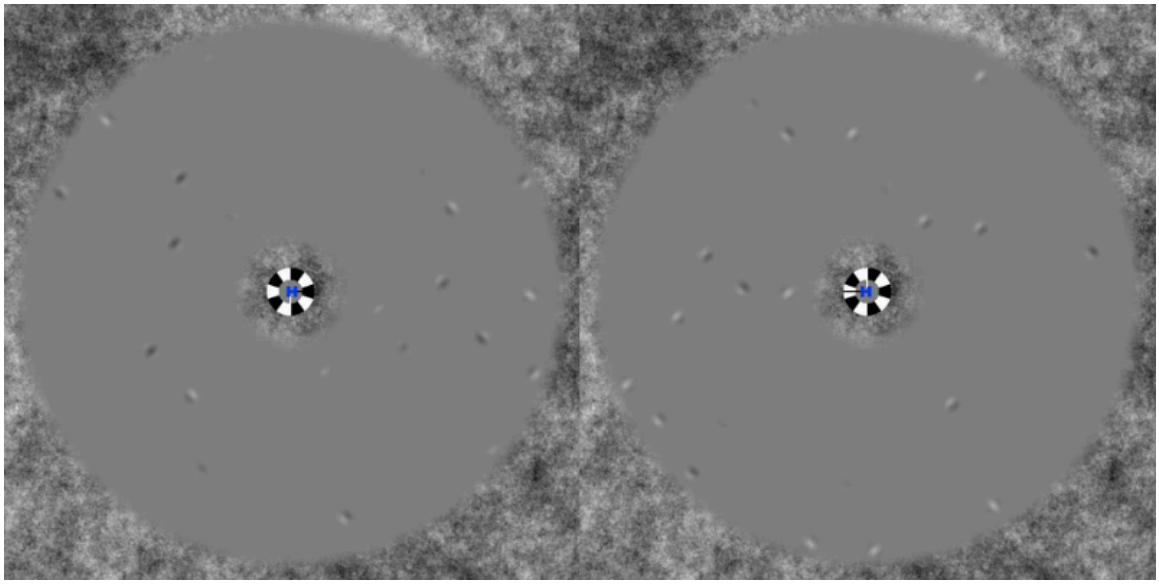

Figure 6.2 Pseudoplaid stimulus movie

A video walkthrough of the pseudoplaid stimulus. Binocular 3D motion percepts can be experienced by cross- or free-fusing the two halves of the display.
Viewable online: http://www.visualstimul.us/pseudoplaids/Walkthrough.html

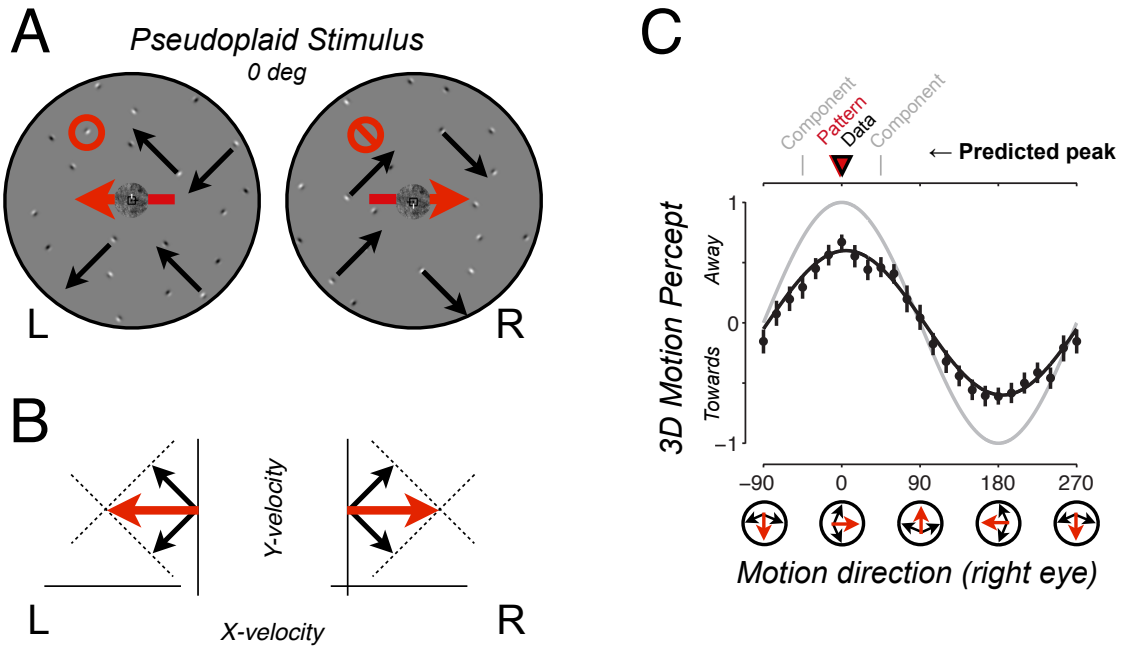

Figure 6.3 Pseudoplaid stimulus schematic & results

(A) An illustration of the pseudoplaid stimulus shows binocularly unmatched arrays of drifting gabors (black arrows & red emphasis) oriented ±45° from the pattern motion direction in each eye. (B) Velocity-space vectors showing the resulting intersection-of-constraints (IOC) pattern motion computation. (C) Perceived 3D motion responses (y-axis; as in Figure 6.1) plotted as a function of the pattern motion direction in the right eye (x-axis; left eye always 180° opposite). As in the plaid condition, the magnitude of reported 3D motion percepts modulated as a function of the horizontal component of pattern motion direction. Adapted from Rokers, Czuba, Cormack, & Huk (2011); see Appendix A for original & complete details.

Additional evidence for a late-stage IOVD computation comes from an adaptation paradigm examining spatial frequency integration in binocular 3D motion processing. Shioiri et al. (2009) used monocular adaptation to drifting gratings to induce a motion aftereffect with components of motion in depth. By exploiting incomplete interocular transfer of monocular MAE (i.e. MAE in unadapted eye is ~50–75% of the adapted eye; Mitchell et al., 1975), an illusory IOVD cue can be induced by means of unequal motion aftereffect magnitudes in the two eyes (Brooks, 2002a; e.g. Fernandez & Farell, 2005). Measurements of spatial frequency selectivity revealed that 3D components of the induced IOVD MAE were robust to changes in the spatial frequency of test stimuli, while the 2D MAE exhibited known spatial frequency selectivity (Cameron, Baker, & Boulton, 1992).

In the interest of specifically addressing binocular motion cues, all of our 3D motion experiments have consisted of stimuli moving in equal and opposite horizontal directions of motion in the two eyes. While this is indeed the most direct case for 3D motion, the trajectories that give rise to opposite directions of retinal motion span only a few degrees of angular motion trajectories that intersect the head (see Figure 2 of Cynader & Regan, 1978). However, under natural viewing conditions the physical extent of rigid real-world objects generate a range of point-trajectories even when moving directly toward the observer. To illustrate a simple example, the schematic depicted in Figure 6.4 (lower left panel) depicts a rigid 20 by 20 cm plane moving at 60 cm/s directly toward the observer (represented by two eyes spaced 6.5 cm apart) from a viewing distance of 70 cm; a close approximation of the experimental stimuli we used to measure 3D speed and eccentricity tuning in Chapter 2. From this stereo viewing geometry it is relatively straightforward to determine the resulting projections produced in the two eyes. Projected monocular velocities are depicted in surface plots in the upper half of figure 6.4. Horizontal and vertical retinal eccentricities are plotted on the x- and y-axes, respectively, absolute magnitude of retinal motion is shown on the z-axis. To emphasize the velocities relevant to 3D motion, surface color corresponds to only the horizontal component of retinal motion. As can be seen in the monocular velocity vectors shown in

the inset of each surface plot, patterns of optic flow could provide monocular cues to 3D motion (see insets; Gibson, 1950). Of course, monocular cues were not present in any of our experimental stimuli, but this example is a useful anchor for understanding how the monocular velocities of rigid 3D motion might constrain the inputs to an IOVD mechanism. The surface in the lower right panel of Figure 6.4 shows the results of an IOVD comparison of horizontal motion components at corresponding locations in the two eyes. For direct comparison with our experimental results, values on the z-axis correspond to one-half the difference in monocular velocities (i.e. the monocular velocity necessary to generate an equivalent IOVD when equal and opposite directions of motion are presented in the two eyes). Surface color represents the angular velocity difference between the two eyes. The velocity vectors from monocular projections are also plotted on the floor of the graph. Note the grey band highlighting the relatively small region where opposite horizontal directions of motion are present in the two eyes. This example reveals not only the utility of IOVD computations under realistic viewing conditions, but also the scope of resulting monocular velocity signals generated from rigid 3D motion.

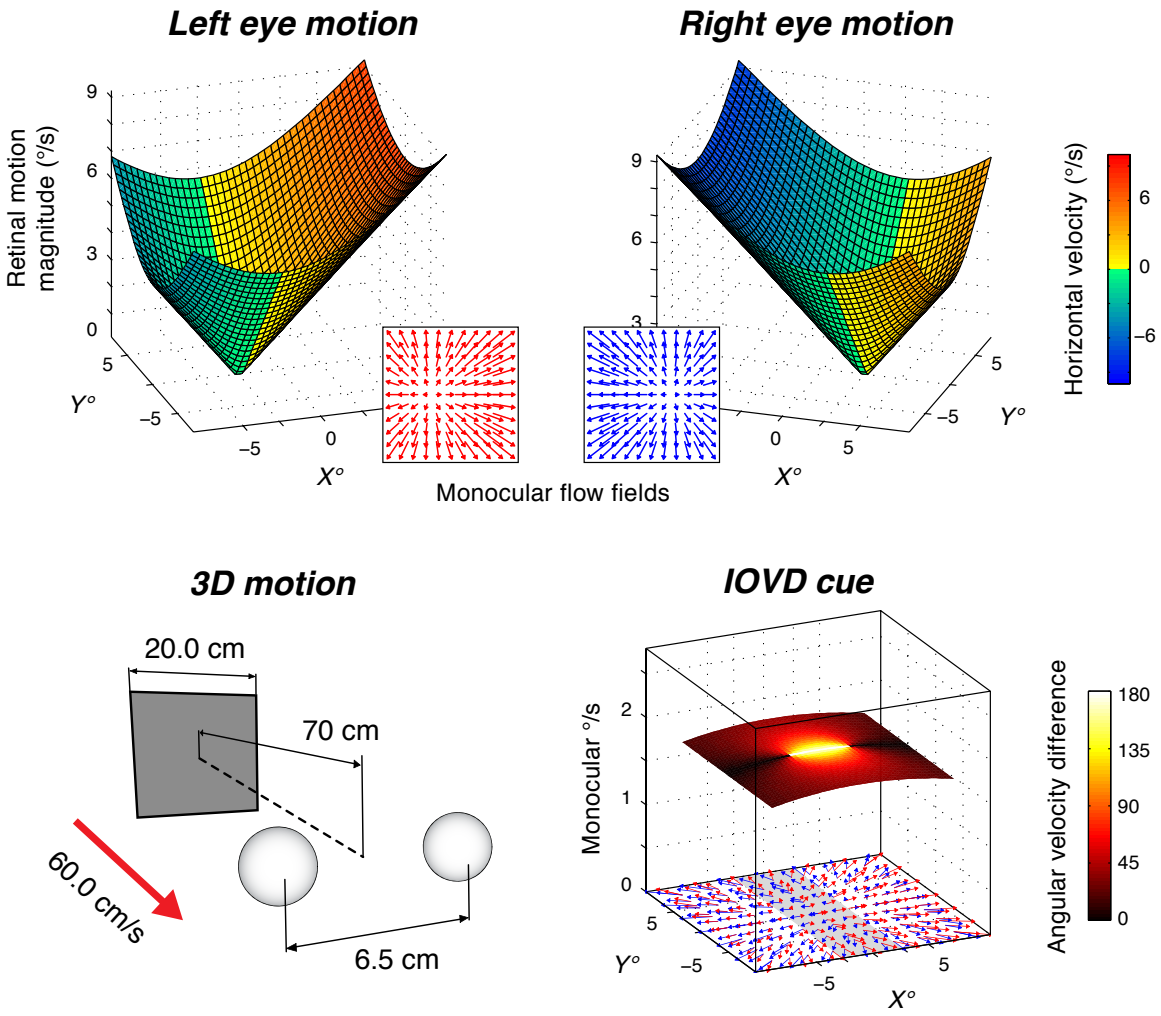

Figure 6.4 Projective geometry of rigid 3D motion

Clockwise from lower left: Schematic of 3D motion of a rigid fronto-parallel plane moving directly toward the middle of the observer's eyes (6.5 cm interpupillary distance). Upper Left & Upper Right: Surface plots of monocular velocity projections in the left and right eye, respectively. The x- and y-axes correspond to projected retinal eccentricity (°), absolute magnitude of retinal velocity (°/s) is shown on the z-axis, and surface color corresponds to the horizontal component of retinal motion in °/s (see color bar). Insets: Velocity vector fields indicating the relative orientation and magnitude of projected monocular velocities across the stimulus area. Lower Right: Surface plot of resulting IOVD cue. Again, x- and y-axes correspond to projected retinal eccentricity (°), z-axis corresponds to one-half the difference in monocular velocities (monocular °/s; i.e. the IOVD equivalent if equal and opposite directions of motion are presented in the two eyes), surface color corresponds to the difference in angular projected monocular velocity in the two eyes. Ground plain (z=0) shows velocity vectors from monocular motion surface insets; grey bar highlights range of projections with opposite horizontal components of motion in the two eyes.

While the understanding of how the binocular signals contribute to 3D motion processing provides insight into how motion signals are combined across the two eyes, new questions are raised about the internal frame of reference for 3D motion processing. Evidenced by the slight curvature of the surface in Figure 6.4 (lower right), an IOVD computation performed directly on the horizontal component of velocity provides information only on the component of motion in the direction of the observer's head. For such a 3D motion signal to be translated into a euclidian, world-centered, frame of reference it must be integrated with 2D retinal motion signals. While electrophysiology, neuroimaging, and psychophysical literature suggest that 2D motion is represented in retinotopic coordinates (Gardner, Merriam, Movshon, & Heeger, 2008; Knapen, Rolfs, & Cavanagh, 2009; Van Essen, Maunsell, & Bixby, 1981), it is unclear just what a 'retinotopic' representation would mean for a signal that is inherently higher dimensional than retinal coordinates allow. The ecological significance of 3D motion may require a representation more closely linked to 3D space. On the other hand, 3D motion information may only be combined with 2D motion when readout by other cortical areas that represent information in body- or object-centered coordinates (Deangelis & Newsome, 2004; Duhamel, Bremmer, BenHamed, & Graf, 1997; van Wezel & Britten, 2002).

CONCLUSIONS & FUTURE DIRECTIONS

The significance of the velocity-based cue to 3D motion perception motivates the amendment of canonical models of motion processing to include the comparison of eye-specific motion signals, and raises new questions about the structure of information represented in cortical areas MT & MST. While the simplest case of a velocity-based 3D motion mechanism would involve only contributions of the horizontal component of monocular motions to the IOVD signal, psychophysical evidence shows that a more complicated mechanism involving monocular pattern motion computations contributes to 3D motion perception (e.g. Rust et al., 2006). Meanwhile, the constellation of results on isolated changing disparity cues suggests that an explicit mechanism for the CD cue may

not contribute to 3D motion per se, but could play a more important role for tracking dynamic disparity information and surface segmentation in locations very near fixation. Current efforts to expand neuroimaging experiments to examine cross-cue adaptation effects in isolated binocular motion cues will provide useful insight to how the distinct characteristics of CD & IOVD cues are integrated to a unified percept of 3D motion.

With mounting evidence for the behavioral significance and characteristic features of IOVD computations, the prospects of uncovering the physiological underpinnings of 3D motion processing have never been better. Important questions remain about the spatial scale IOVD computations and how monocular velocity signals are pooled across corresponding regions of the two eyes. The, at least partial, reliance on integration of monocular pattern motion signals implies a fairly large spatial scale capable of integrating eye specific signals that are not thought to exist at typical stages of pattern motion computation. Further, an IOVD mechanism capable of representing the broad range of 3D trajectories that exist in real world environments would need to compare monocular velocities in a range of directions and speeds appropriate for the location in visual space. In the coming years, I expect renewed interest in the velocity-based cue will fuel significant advancements in our understanding of how the visual system translates two-dimensional monocular inputs into a dynamic representation of the three-dimensional world.

Appendix A: Motion processing with two eyes in three dimensions[9]

INTRODUCTION

The eyes of primates are forward-facing and horizontally offset. As a consequence, an object moving directly toward or away from an observer yields horizontally opposite directions of motion on the two retinae. In addition to the changes in binocular disparity that result from these opposite retinal motions, visual percepts of 3D motion are in part determined by a mechanism that extracts interocular velocity differences (IOVDs) per se. Although there is an increasing body of evidence suggesting a strong and unique contribution of IOVDs (Brooks, 2002a; 2002b; Brooks & Stone, 2004; 2006a; Czuba et al., 2010; Fernandez & Farell, 2005; Rokers, Cormack, & Huk, 2008a; Shioiri et al., 2000; 2008; 2009), it is unclear how motion signals used to make this interocular comparison fit into the known visual hierarchy (Regan & Gray, 2009). We investigated whether the IOVD mechanism operated upon early stages of motion processing typically ascribed to primary visual cortex (V1), or on later stages of processing that extract the motions of patterns, which is typically ascribed to extrastriate areas like the middle temporal area (MT) and related motion processing structures (Khawaja, Tsui, & Pack, 2009; Movshon, Adelson, Gizzi, & Newsome, 1985; Perrone & Thiele, 2002; Simoncelli & Heeger, 1998).

Plaid stimuli have been used to characterize the hierarchical steps involved in 2D motion processing (Adelson & Movshon, 1982; Movshon et al., 1985; Welch, 1989) and 3D depth processing (Delicato & Qian, 2005; Farell, 2003). Superimposing two sinusoidal gratings drifting in different directions produces a plaid pattern that appears to move in a single coherent direction. “Component motion” neurons in V1 respond maximally when either one of the gratings moves in a preferred direction, orthogonal to the cell’s preferred spatial orientation. These neurons have small receptive fields and are

---

[9] This chapter is based on:
Rokers, B., Czuba, T. B., Cormack, L. K., & Huk, A. C. (2011). Motion processing with two eyes in three dimensions. Journal of Vision, 11(2):10, 1–19, doi:10.1167/11.2.10.

thought to function as 1D motion detectors. In contrast, "pattern motion" neurons in areas like MT respond maximally whenever the plaid pattern moves in the cell's preferred direction, regardless of the direction of the individual gratings that constitute the plaid (Movshon et al., 1985; Rodman & Albright, 1989). These pattern motion cells have larger spatial receptive fields than cells in V1 and function as 2D motion detectors.

In addition to the notion that V1 neurons extract component motions, the standard hierarchy of the visual system also dictates that neurons in V1 combine monocular inputs into a single binocular stream (Hubel & Wiesel, 1968), which can thus be thought of as effectively "cyclopean" (Carney & Shadlen, 1993; Shadlen & Carney, 1986). Because interocular velocity differences must (by definition) be extracted prior to complete binocular combination, these two assertions imply that interocular velocity differences must be derived from early component motions, rather than later pattern motion signals.

Contrary to this prediction, in a series of psychophysical experiments described below, we repeatedly found evidence for a strong contribution of eye-specific pattern motions to 3D motion percepts. In the Results section, we describe the dependencies of 3D motion direction discrimination on the relative directions of both component and pattern motions present in dichoptic motion displays. Each experiment in this sequence presents an increasingly specific requirement that the visual system rely on eyespecific pattern motions to compute 3D motion (as distinct from component motions and/or changes in disparity over time). This culminates in a novel "dichoptic pseudoplaid" visual stimulus that lacks conventional binocular matchesV but which still yields 3D motion percepts when it contains global pattern motions that are horizontal and opposite in the two eyes. In the Discussion section, we then review why the overall pattern of results cannot be explained by known disparity-based mechanisms. Together, these results support the conclusion that the brain extracts 3D motion by comparing eye-specific 2D motion signals, using eye-of-origin information widely believed to be unavailable at later stages of visual motion processing.

METHODS

Observers

A total of 7 observers participated in the experiments. All had normal or corrected-to-normal vision. Four observers (the authors, males aged 26–44) were experienced psychophysical observers in both motion and depth experiments. One of the authors (B.R.) and three naive observers participated in a second set of 2AFC experiments. One of the naive observers was an otherwise experienced psychophysical observer. The other 2 naive observers had no history of performing visual psychophysics whatsoever. All naive 3 observers showed similar effects at the individual level, although the more experienced naive observer tended to show larger modulations as a function of our experimental manipulation (yet smaller than the expert author subject). Observers were included based on the criterion that they could easily judge the direction of 3D motion for a large (12.5 deg) grating that drifted in opposite horizontal directions in the two eyes (one potential naive observer was excluded on these grounds).

Experiments were undertaken with the written consent of each observer, and all procedures were approved by The University of Texas at Austin Institutional Review Board. All data were collected at UT Austin, and all observers were recruited from the UT Austin community.

Apparatus

All experiments were performed on a Quad Core Intel Mac Pro with an NVidia GeForce 8800 GT GPU, running Matlab (The Mathworks) and the Psychophysics Toolbox (Brainard, 1997). The stimuli were presented on two 35.0 cm 26.3 cm CRT displays (ViewSonic G90fB, one for each eye; 75 Hz, 1280 1024 pixels) at a viewing distance of 90 cm (21.2–16.3 deg of visual angle). Left- and right-eye half-images were combined using a mirror stereoscope. The luminances of the two displays were linearized using standard gamma-correction procedures, and the mean luminance was 50.6 cd/m2.

Stimuli

All stimuli were presented within an annular aperture (1 deg inner and 6.25 deg outer radii) in a background of 1/f noise (which made it easy for the observers to maintain vergence). Additionally, a small, square fixation mark was placed at the center of the display, which had both horizontal and vertical nonius lines on its perimeter.

The grating and conventional plaid stimuli were spatially apertured (2-dimensional Gaussian, sigma 2.5 deg, centered on the fixation point) and temporally apertured (positive half of a raised cosine, 500 ms half-period, spanning the stimulus duration). Each grating was presented at 10% Michelson (unapertured nominal) contrast, unless otherwise specified. The overall starting (and ending) phase of the drifting gratings was randomized across trials with respect to the Gaussian apertures. On each trial, all gratings started and ended at : binocular phase disparity (i.e., began completely out of phase in the two eyes so that the direction of disparity was ambiguous). No feedback concerning performance was given.

The single-grating stimulus consisted of a single 2 cycles/deg sinusoid drifting at 2 cycles/s in opposite directions in the two eyes. The relatively slow monocular speed (1 deg/s in each eye) corresponds to moderately brisk 3D motion, ranging from 35.2 (toward) to 57.2 (away) cm/s for purely horizontal monocular motions. From trial to trial, grating orientation (and hence, direction) was pseudorandomly selected from a set of directions spanning 0 to 360 deg in 15 deg increments, with 0 defined (by arbitrary convention) as horizontal rightward motion in the right eye (the left eye was always 180 deg opposite in orientation/ direction). The data shown in Figures A.1B and A.5A were collected using a single 10% contrast grating. We also performed this experiment using a 20% contrast grating (thereby equating the net contrast energy to the 2-component plaids used in Experiments 2 and 3) and obtained identical results.

In the plaid experiments, we characterized the 3D direction discrimination performance produced by plaids moving in opposite directions in the two eyes. First, we presented “Type I” plaids consisting of two superimposed drifting gratings in each eye (each with the same parameters described above), with orientations that differed by 120

deg (components oriented ±60 deg relative to the pattern motion direction; Figure A.2A). The contrast values of the two component gratings were added to produce the plaid. The same range of directions and number of trials were employed as in the grating experiment (and in all following experiments).

We also characterized 3D direction discrimination performance produced by "Type II" plaids (Figure A.2C). We again presented two drifting gratings in each eye; this time, however, the components were separated by 30 deg (oriented ±15 deg) and the temporal frequency of one of the gratings was increased to 4 cycles/s. The two components thus had the same spatial frequencies but speeds that differed by a factor of 2. This yields a pattern motion distinct from the individual components, strongly biased toward the "intersection of constraints" ("IOC") direction (Adelson & Movshon, 1982), in this case 51.2 deg away from the mean component direction.

In the pseudoplaid experiments, we characterized 3D direction discrimination performance generated by spatially isolated component motions that specified opposite pattern motions in the two eyes. We constructed the stimulus by presenting 28 small non-overlapping drifting gratings within small stationary Gaussian apertures (standard deviation, 0.1 deg), creating Gabors that were considerably smaller than 1 deg in diameter (i.e., T3 SDs corresponds to a 0.6 deg diameter). Gabors were presented at 20% Michelson contrast (nominal maximum at aperture center), 2 cycles/deg spatial frequency within a 500-ms raised cosine temporal aperture. We presented half (14) of the Gabors in one eye, and the other half in the other. All drifting Gabors were compatible with a single pattern motion direction in each eye, and the pattern motion was in opposite directions between the eyes.

In the 2-component "pseudoplaid" condition (Figure A.3A), each of the Gabors was randomly assigned one of two (±45 deg) orientations (akin to spatially sampling the two components of a 90 deg Type I plaid) and drifted at 2 cycles/s in opposite directions in the two eyes.

In the multi-component "pseudoplaid" condition (Figure A.3C), we oriented the gratings in each aperture randomly, but all component velocities were compatible with a

single (IOC) velocity (this is akin to spatially sampling a complex 2D object with multiple orientations, moving in a single direction). Temporal frequency was 2sinE cycles/s, where E is the angle of the grating relative to the pattern motion direction, so that when the Gabor carrier orientation was exactly perpendicular to the pattern motion direction, the temporal frequency was at its maximum of 2 cycles/s.

Strict constraints were placed on location of individual Gabors within these pseudoplaid stimuli in order to preclude possible binocular (or pattern motion) integration at the scale of conventional V1 mechanisms (Van Essen, Newsome, & Maunsell, 1984). The center of each drifting Gabor was separated by at least 2 deg from the center of any other Gabors in either the contralateral or ipsilateral eye (i.e., throughout the cyclopean view). In combination with the 0.1 deg SD of the Gaussian envelope of each Gabor, and the 8-bit resolution of our displays, this amounted to ~1.4 deg edge-to-edge dichoptic and monocular spacing, as the rendered contrast of each Gabor was effectively zero at 3 SDs (i.e., 0.3 deg).

Although the 2AFC task is in very common use, we originally used the 5-point rating scale task because it allowed us to see systematic effects of direction with stimulus strengths well above conventional psychophysical threshold (e.g., $d' \gg 2$). Likewise, we had originally focused on expert observers because of the large number of experimental conditions, coupled with our desire to make rather fine-grained quantitative inferences about the location of peaks. To assess performance in naive observers and using a 2AFC task, we performed additional measurements with two issues in mind. First, we expected that the naive (non-expert) observers might be generally less sensitive (and perhaps more variable) than the expert observers: we therefore collected data from one expert observer as a reference. Second, we needed to modify some stimulus parameters to bring 2AFC performance closer to threshold (thereby avoiding the ceiling effects that may be bypassed in the rating scale task): thus for conventional grating and plaid stimuli (e.g., gratings, type I plaids, and type II plaids), presentation time was reduced to 250 ms. Accordingly, temporal frequency was doubled (to 4 cycles/s from 2 cycles/s) to maintain whole temporal cycles during presentation.

In each experimental condition, all drifting components could thus start and end every trial at a half-cycle of binocular phase disparity, drifting for exactly one or two cycles during each trial. Each trial thus began and ended at an ambiguous disparity (owing to the periodic nature of the stimuli), which could be perceived as either near or far. Instantaneous starting, ending, as well as average disparities therefore could not serve as potential cues for task performance.

Procedure and task

The stereoscope was initially adjusted so that the vergence demand was appropriate for the viewing distance given a typical interocular distance. Prior to each session, each observer made further minor adjustments so that the nonius markers were aligned both horizontally and vertically, and vergence was comfortable. Observers were instructed to maintain fixation for the duration of each experimental session, and the four expert observers were all experienced in monitoring vergence during 3D motion psychophysics, minimizing the possibility of unintended binocular matches.

Before participating in the experiments, the naive observers were introduced to demonstration versions of the stimuli and task, viewing them using active stereo shutter glasses (NVIDIA) and a 120-Hz DLP projector (DepthQ). The observers were instructed to try to discriminate the global 3D direction (toward/away) of pseudoplaid stimuli, ignoring the frontoparallel motions of the individual Gabor elements. After these preliminary demonstrations, the naive observers were then assisted in adjusting the stereoscope in the main apparatus, and they did not report any discomfort or problems with fusion of the surround.

For the expert observers, on each trial, the stimulus was presented for a single 500-ms interval, and the observer responded via a key press. For each of the 4 observers, 30 repetitions of each of 24 directions were pseudorandomly distributed across 1 or 2 runs. Observers performed a signal detection 5-point confidence rating direction discrimination task (toward or away; the 5 potential responses corresponded to high confidence away, low confidence away, neutral/ambiguous, low confidence toward, high

confidence toward). This can also be thought of as a 2AFC task, where the response on each trial is assigned one of three confidence levels (high, low, totally ambiguous). The main advantage of this technique is that it simultaneously captures multiple criteria and can thus be used to quickly generate an ROC curve (see Figure A.5B), hence separating bias and sensitivity.

We acquired similar results with 3 naive observers (and one of the authors) using a simpler 2-alternative forced choice (2AFC) paradigm, using stimuli closer to 3D motion detection threshold. The temporal parameters were changed to a shorter 250-ms presentation at a faster 4 cycles/s. Some ceiling effects are evident for the experienced observer in the 2AFC data, but even with saturated accuracy levels that distort the sinusoidal dependence on direction, the location of the peaks of his curve were still consistent with all the other results. These ceiling effects were not a factor for the expert observers, since they reported motion direction and associated confidence on a 5-point scale, rather than making a 2AFC. Indeed, we incorporated confidence ratings to avoid such ceiling effects for the most effective stimulus conditions (Figure A.5B demonstrates that the results from the signal detection confidence rating task continued to follow systematic modulations for d′ values well above 3). We note that both 2AFC and multiple-level rating scale tasks can be analyzed to extract standard quantities like receiver operating characteristic (ROC) curves and d′ (Green & Swets, 1966; Macmillan & Creelman, 2005).

Data analysis

For the expert observers, we encoded each observer's 5-point responses on a linear rating scale from 1 (away, high confidence) to –1 (toward, high confidence); intermediate responses of 0.5 indicated low confidence away and –0.5 indicated low confidence toward, and a 0 rating indicated zero toward/away confidence. We plotted the mean rating as a function of direction (defining horizontal right-eye motion as "0 deg"). All statistics were estimated using a bootstrap procedure, resampling the data from each run and from each subject 1000 times; this propagates inter-run and inter-subject

variability into all estimates of variance. Error bars on each data point represent 95% confidence intervals, equivalent to 2 standard errors of the mean. These 3D motion perceptual tuning curves are shown in Figures A.1B, A.2B, A.2D, A.3C, and A.3F. (We also derive ROC curves and d′ for these data, as shown in Figure A.5B.)

To quantify the dependence of perceived 3D motion on direction, we fit the perceptual tuning curves (i.e., mean rating as a function of direction) with a sinusoidal function (with amplitude and phase as free parameters), minimizing RMS error (black curves, Figures A.1B, A.2B, A.2D, A.3C, and A.3F). The fitted amplitude specifies the height at the peak of the curve, and the fitted (cosine) phase indicates the location of the peak of the curve. We estimated 95% confidence intervals on these fitted parameters using a bootstrap (1000 iterations). We then compared the phase confidence interval (i.e., the confidence interval about the estimated location of the peak of the curve) with relevant component and pattern motion direction predictions for each experiment. For example, we could test whether the confidence interval about the fitted peak contained the expected peak location based on horizontal pattern motions, as well as whether it excluded the predicted peak based on alternatives (such as the vector average of the component motions).

For the naive observers who performed a 2AFC direction discrimination task (toward versus away), we simply measured the proportion of "away" judgments as a function of direction, and then performed the same analyses as for the rating scale data.

As expected, the naive observers were not as skilled in adjusting the mirror stereoscope to be in perfect alignment (as compared to the expert observers). Although they confirmed good binocular fusion and maintenance of vergence, it was likely that slight optical misalignments (i.e., mixtures of pan and tilt) produced images that were slightly rotated relative to ground truth (i.e., the "0 deg" condition was actually not perfectly vertical). We estimated this misalignment for each observer using their responses in the drifting grating experiment. Using the peak of this baseline condition as the corrected "0 deg," we then subtracted this value to shift the curves in subsequent

conditions. This correction was only on the order of a few degrees at most, and we note that it does not affect the pattern of results across conditions.

Rating scale tasks, such as the one used in our expert observer experiments, are well suited to conventional signal detection receiver operatic characteristic analysis. For the ROC analysis (Figure A.5B), we used standard n-afc rating scale procedures (Green & Swets, 1966). Each response boundary (e.g., "G0" vs. "Q0") reflects a criterion, thus yielding 4 criteria from our 5-point scale. We then tallied the proportion of hits and false alarms under a specific criterion for a given absolute stimulus direction (which was hypothesized to correspond to horizontal IOVD signal strength). This resulted in 4 data points per ROC curve. The results of this analysis are shown in Figure A.5B for the single-grating data to illustrate the utility of the rating scale method and its relationship to 2AFC.

In an additional control experiment, we verified that expert observers did not rely on joint information on eye of origin of stimulus elements and monocular direction of motion. The observers performed a 2AFC direction discrimination task on a dichoptic multiple-grating stimulus. This stimulus was similar to a dichoptic pseudoplaid, except all stimulus elements had the same orientation within each monocular view (as before, they moved in opposite directions in the two eyes; Figure A.4A). Observers judged 90 trials each in two motion conditions (vertical motion or horizontal motion). For the vertical motion, observers judged whether the motion in the right eye was upward or downward (as an assay of their ability to perform the posited joint eye-of-origin/ direction-of-motion task); for the horizontal motion, they judged whether the 3D motion was "toward" or "away" (as an assay of their ability to discriminate 3D motion direction). We computed the proportion of times observers reported the stimulus as moving away (horizontal condition) or upward in the right eye (for the vertical condition).

RESULTS

We performed a series of experiments in which observers judged 3D motion direction as a function of the monocular directions of either component (1D) or pattern

(2D) motions. The initial baseline experiment, using a single 1D grating in each eye, serves to explain the general methods and logic of the approach. The following experiment, using a Type I plaid in each eye, extends this logic to 2D dichoptic motions. Then, the remaining experiments (Type II plaids and dichoptic pseudoplaids) represent the key results that characterize the eye-specific motion signals used to estimate 3D motion.

Baseline Experiment: Dichoptic Gratings

We first characterized 3D motion direction discrimination performance as observers viewed a single grating that drifted in opposite directions in the two eyes (Figure A.1A). These single-grating data served as a baseline for the rest of the experiments that employed multiple component motions in each eye. We varied grating orientation from trial to trial, and on each trial, the expert observer judged the direction of perceived 3D motion (indicating either motion toward or away, at one of three levels of confidence in a signal detection rating scale task; Figure A.1B; throughout the following experiments, we plot the mean rating and refer to this quantity as "direction discrimination performance"). On each trial, the right and left eyes' gratings always moved in the polar opposite direction from one another (i.e., 180 deg apart).

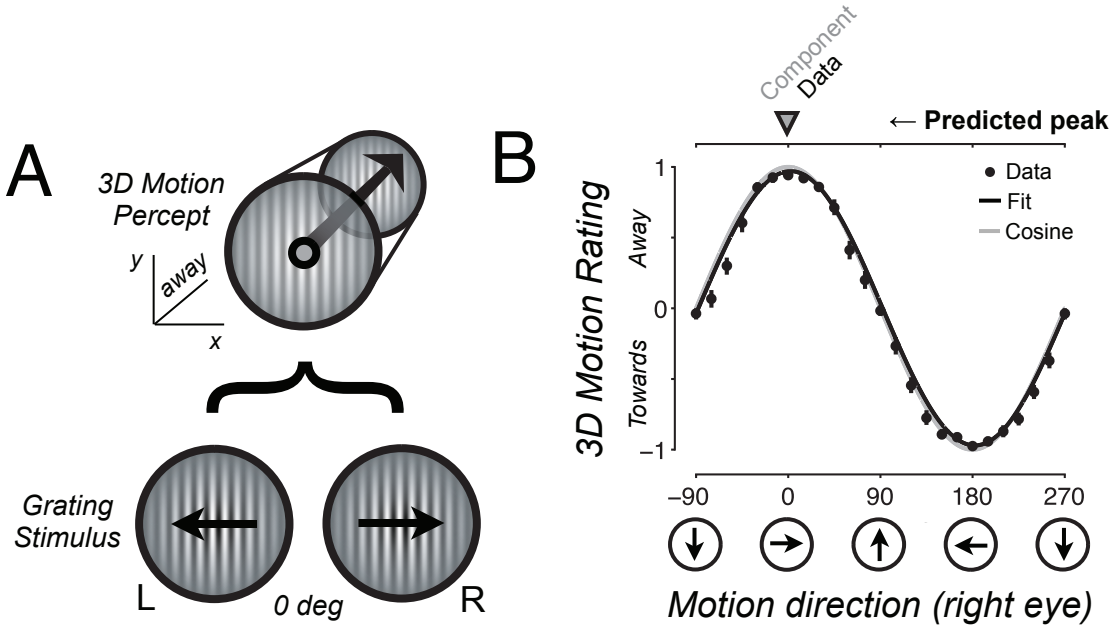

Figure A.1 Dichoptic grating stimulus schematic & results

Three-dimensional motion direction discrimination when viewing oppositely moving gratings in the two eyes. (A) Illustration of the 3D motion percept and dichoptic grating stimulus. All observers viewed gratings drifting in opposite directions in the two eyes (L, left eye; R, right eye). Expert observers performed a rating scale 3D direction discrimination task as grating orientation (and hence direction) was varied from trial to trial ("away" or "toward", at either high, low, or zero confidence). Gratings always moved in opposite directions in the 2 eyes. By convention, we labeled the condition illustrated here (rightward horizontal motion in the right eye) as "0 deg." Trials with other directions of motion were created by rotating both eyes' gratings in 15 deg increments. This maintained a matched stimulus orientation in both eyes on each trial, while also yielding 180 deg opposite directions of motion. (B) Three-dimensional motion direction discrimination resulting from viewing a grating moving in opposite direction in the two eyes. Y-axis plots the mean 3D motion rating (signed direction and confidence, in a signal detection rating scale task; see Figure A.5B for illustration of ROC curves and d′ estimates from this rating scale data). X-axis denotes the direction of grating motion in the right eye (graphically indicated by the icons below the plot; left eye motion was always 180 deg opposite). The stimulus supported the clearest 3D direction discrimination (away or toward) when the grating drifted horizontally (i.e., at 0 and 180 deg, respectively) but not when it drifted vertically (90, 270 deg). Black curve indicates a sinusoidal fit to the data. This curve closely matches a reference unit-amplitude cosine function with a peak at 0 deg (gray line). Each data point (black dot) indicates the mean response combined across subjects. The sinusoidal relationship between monocular motion direction and 3D motion rating indicates that the 3D motion discrimination

performance was proportional to the horizontal portion of the monocular motions. In this and all following figures, similar patterns were observed in each individual. Error bars depict 95% confidence intervals on each point and are sometimes smaller than the plotting symbols. The gray solid triangle above the upper x-axis indicates right eye motion direction for which the 3D motion direction discrimination performance (of "away" motion) was predicted to be highest based on the horizontal fraction of grating motion (0 deg). The open black triangle indicates the best fit peak from the data (2.2 deg).

Three-dimensional motion direction discrimination performance was highest (i.e., correctly judged as "away" or "toward" with the highest levels of confidence) when the grating drifted horizontally (0, 180 deg), intermediate at intermediate directions, and observers were unable to reliably discriminate 3D motion direction when the grating's motion was completely vertical (90, 270 deg). This perceptual tuning curve was well characterized by a sinusoidal function, which we use throughout the following experiments as a descriptive fit that allows us to estimate the location of the perceptual sensitivity peak with respect to monocular direction. In this experiment with a single grating in each eye, the 95% confidence interval on the location of this peak spanned [0, 3] deg. The apparent cosine phase suggests that 3D motion sensitivity was proportional to the horizontal portion of the 2D grating motion vectors.

Dichoptic type I plaid experiment

Introducing the idea of eye-specific pattern motions

We then measured the dependence of 3D motion direction discrimination as observers viewed conventional drifting plaids in each eye. In the Type I plaid condition, both eyes' plaids were composed of two superimposed gratings rotated by ±60 deg (Figure A.2A). Figure A.2B shows observers' 3D motion direction discrimination, plotted as a function of the plaid pattern motion direction (in the right eye). Direction discrimination performance peaked when the pattern direction was horizontal (0 and 180 deg; [–2, 1] deg, 95% CI on peak location)–such that component grating directions were at ±60 deg from horizontal. In contrast, when either one of the component grating motions was horizontal (i.e., when the pattern motion was at -60, 60, 120, or 240 deg), direction discrimination performance was low. The fact that the perceptual tuning curve was unimodal, instead of containing peaks at ±60 deg, indicates that discrimination performance did not derive from a winner-take-all mechanism that operated solely upon the most horizontal component motion. Of course, the sum of two sinusoid response curves with phases at ±60 deg is simply a sinusoid of the same frequency, with a phase at 0 deg. Thus, a mechanism that perfectly summed the horizontal components of the

disparities—or the velocities for that matter—would produce sinusoidal response curves centered at 0 deg. The additional experiments that follow rule out a wide set of disparity and component motion-based accounts of these initial results.

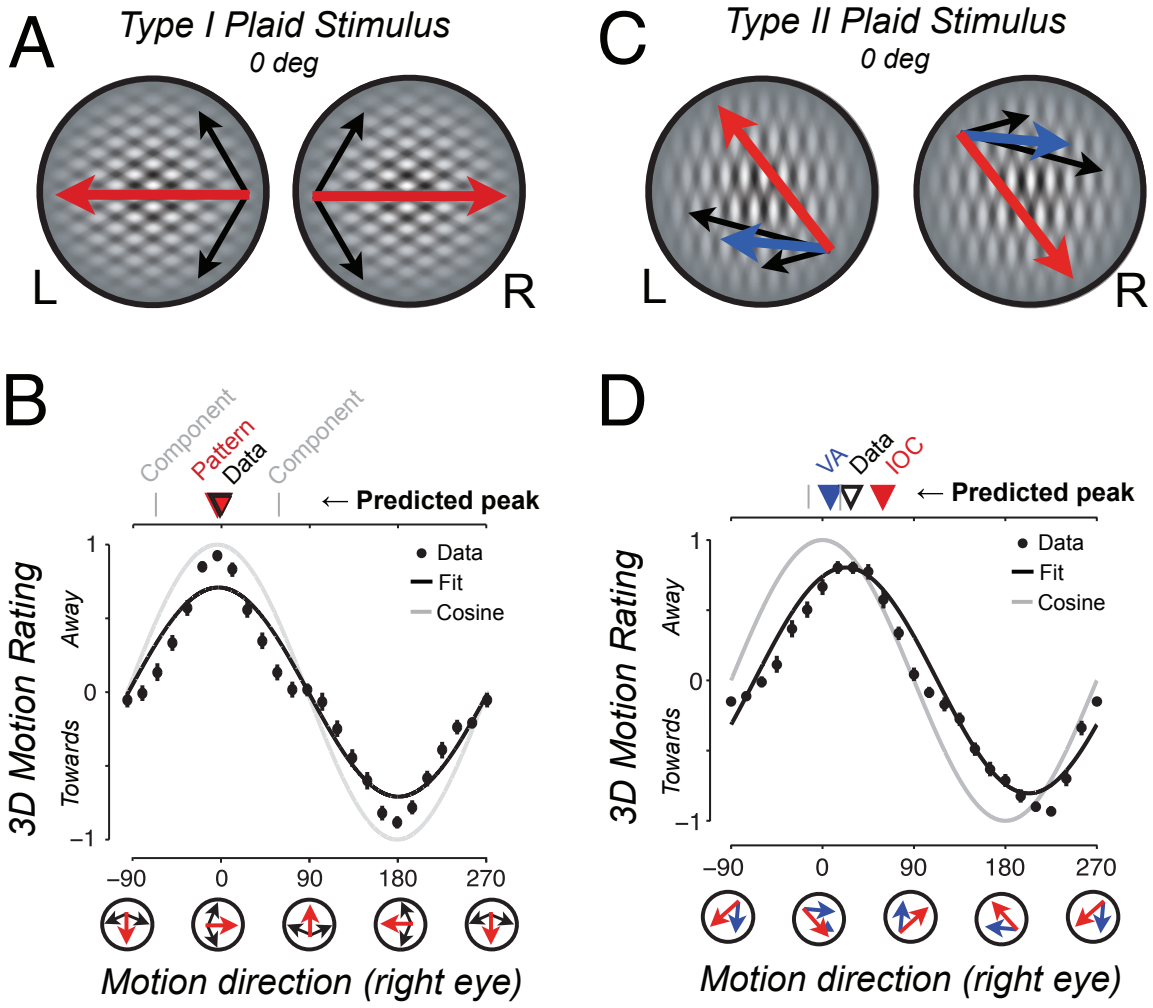

Figure A.2 Dichoptic Type-I & Type-II stimulus schematics & results

Three-dimensional motion sensitivity when viewing dichoptic plaids with opposite pattern motions in the two eyes. (A) Illustration of the Type I plaid stimulus. Observers viewed two superimposed gratings separated by 120 deg (black arrows) forming a Type I plaid (pattern motion direction, red arrows) moving in opposite direction in each eye (L, left; R, right). The condition illustrated here corresponds to the 0 deg condition in lower panels (i.e., rightward pattern motion in the right eye; left eye motions were always 180 deg opposite). (B) Three-dimensional motion direction discrimination performance when viewing a Type I plaid moving in opposite direction in the two eyes. Format similar to Figure A.1B; 3D motion direction discrimination performance (y-axis) as a function of pattern motion direction in the right eye (x-axis). Three-dimensional motion performance (mean rating; as in Figure A.1B) was highest when the pattern motion was horizontal (0, 180 deg) and very weak when either of the components moved horizontally (–60, 60, 120, 240 deg). Black curve indicates a best-fit sinusoid; gray curve indicates a reference unit-amplitude cosine function identical to Figure A.1B. Icons below the plot indicate the directions of right eye motion. Symbols above the upper x-axis indicate right eye motion

directions that would predict the highest 3D performance (for “away” motion), based on component motion (±60 deg; gray ticks), pattern motion (0 deg; red solid triangle), and best fit to the data (–0.5 deg; open black triangle), suggesting a dependence of the 3D percept on the plaid pattern motion direction. (C) Illustration of the Type II plaid stimulus. Observers viewed a Type II plaid moving in opposite direction in the two eyes. Component motions associated with the two superimposed gratings are indicated with black arrows; pattern motion of the plaid (by IOC) is indicated by the red arrows; vector average of the components is indicated by the blue arrows. The condition illustrated here corresponds to the “0 deg” point in the next panel (i.e., 0 deg average component direction in the right eye; left eye motions were always opposite). Temporal frequency of the faster component was twice that of the slower component. (D) Three-dimensional motion direction discrimination performance when viewing oppositely moving Type II plaids in the two eyes. X-axis indicates the average direction of the two components (i.e., 0 deg indicates the stimulus schematized in (A) above). Three-dimensional motion discrimination performance varied as a function of direction. The fitted peak of the perceptual tuning curve fell at 23.0 deg (black open triangle above plot; bounding 95% confidence interval, [20 25] deg), which is well past the vector average direction (5.1 deg; blue triangle), outside the range of directions spanned by the individual components (±15 deg; gray ticks), and shifted toward the IOC direction (51.2 deg; red solid triangle). Thus, the amount of shift cannot be explained by a mechanism based on individual component contributions (or on their vector average) alone and indicates a contribution of pattern motion. Icons below the plot depict the directions of right eye motion.

Dichoptic type II plaid experiment

A first test for a unique contribution of eye-specific pattern motions

To more directly pit pattern motion against component motion (and corresponding temporally changing horizontal disparities), we presented drifting Type II plaids (Adelson & Movshon, 1982) in each eye (Figure A.2C). In these plaids, the pattern motion falls outside the range of the two component motions and is not the simple result of their vector average (or sum). We generated Type II plaids by superimposing a pair of component gratings with a small angle between them (±15 deg), moving at different velocities (1 and 2 deg/s). Under our viewing conditions (and consistent with prior work; Adelson & Movshon, 1982; Welch, 1989; Yo & Wilson, 1992), the resulting pattern motion direction in each eye was well approximated by the "intersection of constraints" (IOC) direction—the single 2D velocity geometrically consistent with both component motions.

Figure A.2D shows 3D motion direction discrimination performance, plotted as a function of the mean component motion direction (right eye; left eye motions were 180 deg opposite). If 3D motion direction discrimination depended solely on the interocular comparison of the vector average (or sum) of component motion signals, the peaks of the tuning curve would fall near 0 deg (VA, blue triangle; 5.1 deg) and certainly within the ±15 deg range of directions spanned by the component motions (vertical gray tick marks). Instead, the peak of the curve (open black triangle; [20, 25] deg, 95% CI) was considerably shifted toward the IOC direction (red triangle; 51.2 deg), reliably outside the component motion range (±15 deg). Because the shift toward IOC was significant but not complete, it is certainly possible that 3D motion percepts reflected a mixture of 2D pattern and 1D component motions.

So far, the patterns of 3D motion direction discrimination resulting from dichoptic Type I and Type II plaids suggest a parsimonious explanation based on a mechanism that computes IOVDs using eye-specific pattern motions: the experiment that follows introduces a less conventional stimulus that compellingly supports this explanation.

Although we observed reliable direction discrimination performance in the absence of feedback (consistent with observers' subjective reports of solid cyclopean percepts for horizontal and near-horizontal retinal motions), binocular rivalry can certainly occur in dichoptic viewing (L. Liu, Tyler, & Schor, 1992) and might be related to the weaker direction discrimination performance we measured as monocular motions approached vertical.

Although these results require some sort of 2D computation prior to the 3D motion mechanism, there is still the possibility that some of these results depend upon 2D disparity matches, rather than 2D velocities. Some evidence exists for two-dimensional binocular matching, instead of simple 1D horizontal disparity extraction (Farell, Chai, & Fernandez, 2009). Although the primacy and relevance of such 2D mechanisms is not yet clear, in principle such 2D disparity processing could complicate interpretations based on (the geometrically equivalent) interocular pattern motions. We therefore sought to design a stimulus that retained opposite pattern motions in the two eyes but that avoided binocular overlap of constituent elements to sidestep rivalry and/or 2D disparity processing.

Dichoptic pseudoplaid experiments

Evidence for eye-specific pattern motions

To perform a more stringent test for IOVDs based on eye-specific pattern motions, we capitalized on the fact that the spatial scale of extrastriate pattern motion mechanisms in primates is significantly larger than that of component motion mechanisms in V1 (Movshon et al., 1985; Van Essen et al., 1984). The resulting "dichoptic pseudoplaid" stimulus—shown in Figure A.3A—contains opposite pattern motion signals in the two eyes that exceeded the spatial scale of component motion mechanisms. Scattered Gabor elements (i.e. individual, drifting sinusoidal gratings within small, stationary Gaussian apertures; Amano, Edwards, Badcock, & Nishida, 2009; Majaj, Carandini, & Movshon, 2007) were arranged to produce a global pattern motion signal in each eye that could only be recovered by integration over multiple elements

across space (Clark & bradley, 2008; Watamaniuk et al., 1989; Watamaniuk & Sekuler, 1992). Crucially, the elements in one eye's view were not matched binocularly—a Gabor in one eye was always farther than 1.4 deg (edge to edge, where "edge" indicates the location at which our Gabors fell to the background luminance) from any other Gabor (in either the same or the other eye). This spacing ensured that no significant processing beyond the extraction of monocular component motion could occur in single units within primary visual cortex given known measures of primate V1 receptive fields at these eccentricities (Van Essen et al., 1984). Similar to the conventional plaid stimuli, each eye's stimulus contained 2 distinct (±45 deg) component motions, producing a pattern motion different from that of the individual Gabor elements (schematized in Figure A. 3B). In addition, as in the earlier grating and plaid experiments, pattern motion direction was always 180 deg opposite in the two eyes.

Observers again discriminated 3D motion direction as we varied the pattern motion direction (Figure A.3C). Three-dimensional motion direction discrimination performance was highest when the global pattern motion in each eye was horizontal ([0, 9] deg, 95% CI on peak location). When pattern motion in each eye was vertical, observers were no longer able to discriminate 3D motion direction. Not surprisingly, discrimination performance was reduced relative to those observed using the binocularly paired conventional plaids, given that this dichoptic pseudoplaid stimulus consisted only of a small number of sparse, monocularly visible Gabor elements. What is surprising is that observers perceived 3D motion at all—instead of a constellation of randomly oriented monocular (or rivalrous) Gabors. This demonstrates that an interocular comparison of pattern motions (in the absence of conventional binocular disparities) is sufficient to yield percepts of 3D motion.

The shape of the perceptual tuning curve argues against the observers' reliance on unintended binocular overlap of elements (perhaps due to vergence eye movements or unexpectedly large but finely tuned orientation receptive fields). Had this been the case, percepts of 3D motion should be strongest when one of the components moved horizontally—but performance under such conditions (±45 deg) was in fact relatively low

(Figure A.3C). Moreover, the stimulus had a 1/f central and peripheral textured background, which provided a powerful anchor for vergence across a range of visual scales.

As a final challenge to interocular pattern motion mechanisms, we further generalized the stimulus so that each drifting Gabor had an independent, random orientation uniformly selected over all 180 deg of possible orientations (Figure A.3D). We then assigned each Gabor a velocity consistent with a single, global, pattern motion velocity (Amano et al., 2009; Majaj et al., 2007)—opposite in the two eyes (Figure A. 3E). This "multi-component pseudoplaid" simulates a single moving object with different dominant orientations at different spatial locations (such as a zebra), viewed dichoptically through apertures that differentially occlude the 2 eyes' views (such as foliage close to one's face). As in the previous experiment, 3D motion direction discrimination performance peaked when the pattern motions were horizontal and opposite (Figure A. 3F; [–9, 0] deg, 95% CI), suggesting that the randomly oriented component motion signals are combined into a single pattern motion signal for each eye, prior to the extraction of 3D motion. As with the conventional gratings and plaids, the patterns of results with dichoptic pseudoplaids were confirmed in naive subjects and using a 2AFC direction discrimination task (Figures A.6C and A.6D).

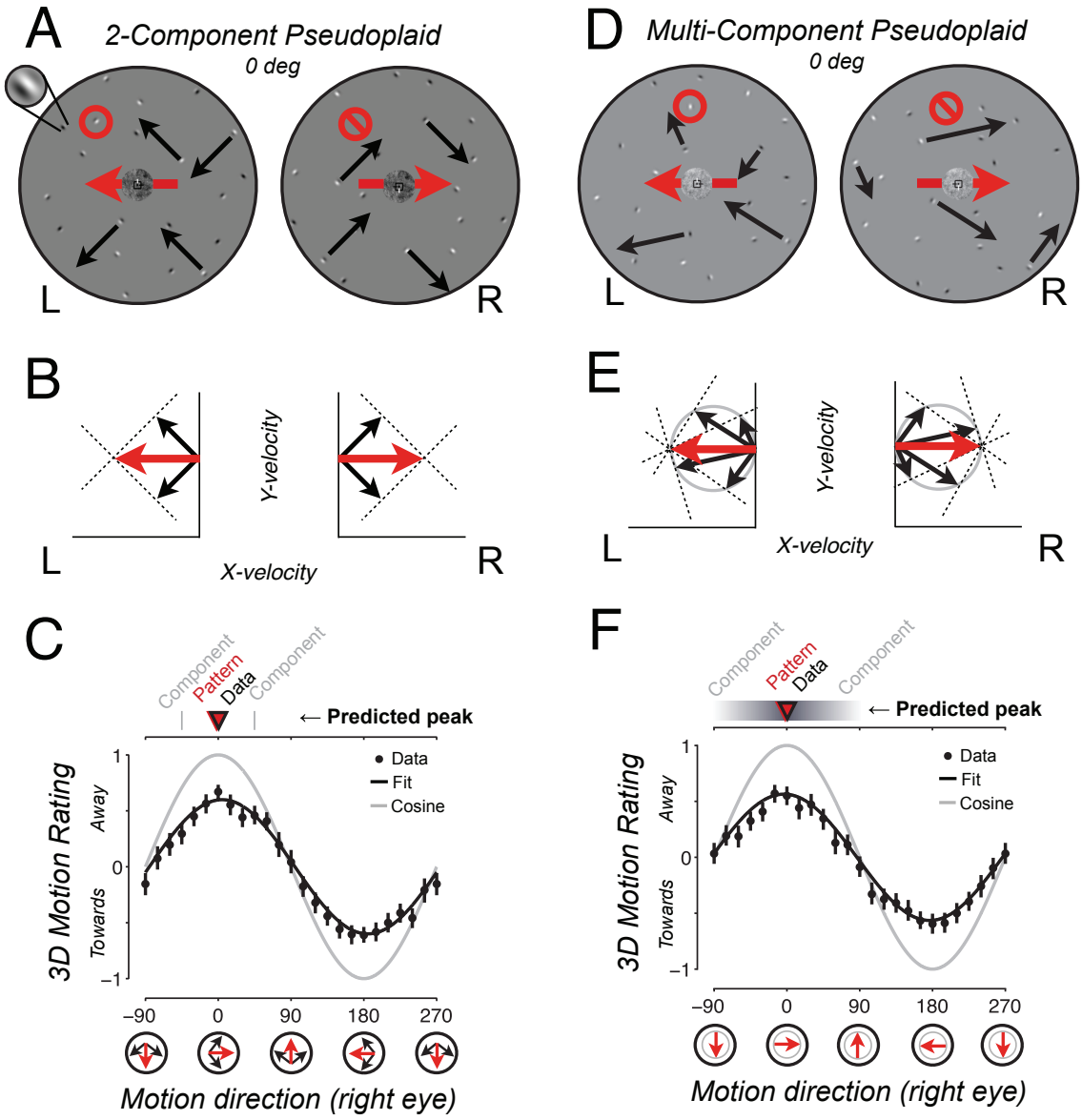

Figure A.3 Two- and multi-component pseudoplaid stimulus schematics & results

Three-dimensional motion sensitivity when viewing dichoptic pseudoplaids with opposite pattern motions in the two eyes. (A) Illustration of the 2-component pseudoplaid stimulus. Observers viewed fields of spatially separated Gabor elements in the left (L) and right (R) eyes. In each eye's pseudoplaid, Gabors were oriented at 90 deg relative to one another (i.e., akin to a ±45 deg Type I plaid but with the components represented as spatially separated Gabor elements instead of overlapping gratings) and randomly distributed in space. Inset top left shows a magnified view of a single Gabor element. Critically, Gabors in the left eye (i.e., red circle) were separated by at least 1.4 deg (edge to edge) from any Gabors in the corresponding right eye's half-image (crossed red circle; see Methods section for more details). Gabors in the left eye's half-image drifted in opposite directions to those in the right eye (black arrows, only some arrows shown for

clarity, in the actual stimulus all Gabors drifted within their stationary Gaussian envelope). The condition illustrated here corresponds to the 0 deg condition in lower panels (i.e., rightward global pattern motion (red arrow) in the right eye). (B) Velocity-space representation of the 2-component pseudoplaid stimulus. Velocity vectors for the left and right eyes' views of the 0 deg pseudoplaid stimulus. Horizontal axis, horizontal velocities; vertical axis, vertical velocities. Black arrows, component motions, representing the Gabor component arrows overlaid on the stimulus illustration shown in (A) in velocity space. Dashed lines depict 1D motion constraint lines; red arrows, global pattern motion. (C) Three-dimensional motion direction discrimination when viewing a 2-component pseudoplaid drifting in opposite direction in the two eyes. X-axis indicates the global pattern direction (i.e., 0 deg corresponds to the stimulus shown in (A) above). Three-dimensional motion direction discrimination performance varied as a function of global pattern direction, with a peak near 0 deg (95% confidence interval on fitted peak location, [0, 9] deg). Three-dimensional motion discrimination performance was low when either pseudoplaid component drifted horizontally (–45, 45, 135, 225 deg). Instead, performance peaked when the pattern motion was horizontal (0, 180 deg), demonstrating a primary dependence on interocular pattern motion and not local component motions. Icons below the plot indicate the directions of right eye motion. Symbols above the upper x-axis depict key elements: Right eye motion direction predicting highest performance (i.e., discrimination of "away" motion) based on either component motion (±45 deg; gray ticks) or pattern motion (0 deg; red solid triangle); best fit to the data (4.6 deg; open black triangle). (D) Illustration of the multi-component pseudoplaid stimulus. Observers viewed a stimulus identical to the one shown in (A), except that the orientations of all Gabors were fully randomized (i.e., uniformly distributed throughout all possible orientations) while their individual speeds were tailored to be consistent with a single pattern motion velocity (by IOC). We call this version of the dichoptic pseudoplaid stimulus "multi-component" simply because it contains multiple compatible component motions. Stimulus depicted here corresponds to the 0 deg condition (i.e., rightward pattern motion in the right eye; left eye was always opposite). For clarity, motions of only some of the Gabors are indicated (black arrows); all elements drifted in the actual stimulus. (E) Velocity-space representation of the multi-component pseudoplaid stimulus. Similar format to (B). Black arrows indicate various component motions, corresponding to a range of orientations as depicted in (D). In this figure, only some of the arrows are shown for clarity, the actual stimulus specified 14 component motions in each eye, drawn randomly from a uniform distribution. Dashed lines depict 1D component motion constraint lines. Red arrow indicates global pattern motion as obtained by intersection of constraints. To be consistent with a single pattern motion, all component velocities produced by the randomly oriented Gabors have to fall on a circle in velocity space (gray circle). (F) Three-dimensional motion direction discrimination when viewing the multi-component pseudoplaid stimulus. Format similar to (C). X-axis depicts the global pattern motion direction (i.e., 0 deg corresponds to the stimulus shown in (A) above). Performance (y-axis) varied as a function of global pattern direction, with a peak near 0 deg (95% confidence interval on fitted peak location, [–9, 0] deg). Icons below the plot

depict the directions of right eye motion. Symbols above the upper x-axis indicate key elements: Right eye direction predicting highest performance (i.e., discrimination of "away" motion) based on either component motion (continuous gray band) or pattern motion (0 deg; red solid triangle); best fit to the data (–3.5 deg; open black triangle).

Eye-of-origin & monocular motion control experiments

We conducted a control experiment to explicitly test whether observers might have performed the 3D motion task by covertly performing a joint eye-of-origin and direction-of-motion discrimination—essentially knowing which eye saw what direction of 2D motion, and then mapping that correctly on to a 3D direction response despite the absence of any feedback (see Figure A.4A). In separate runs, we had the four expert observers view the multi-component dichoptic pseudoplaids when all monocular motions were either purely vertical (and opposite in the two eyes) or horizontal (and opposite in the two eyes). When the motions were vertical, observers attempted to perform a joint eye-of-origin and direction-of-motion task (i.e., "which eye saw upward motion?"). When the motions were horizontal, observers performed a standard 2AFC 3D motion direction discrimination ("toward or away?"). Accuracy on the vertical task tested observers' ability to jointly discriminate monocular direction and eye of origin—and performance in all observers was close to (and often statistically indistinguishable from) chance. Accuracy on the horizontal task was near perfect—far higher than would be predicted from the vertical eye-specific direction-of-motion performance—leaving little possibility that observers based their responses on anything other than the perceived direction of 3D motion. If observers had relied on a cognitive strategy or had simply closed one eye, they should have performed equally well in both conditions. Although one logical possibility remains—that observers can perform such a joint eye-of-origin and direction-of-motion task for horizontal motions (such as in the main experiments) but not for vertical motions (as in this control)—we think this is not likely enough to warrant further consideration unless future data support such an odd proposition.

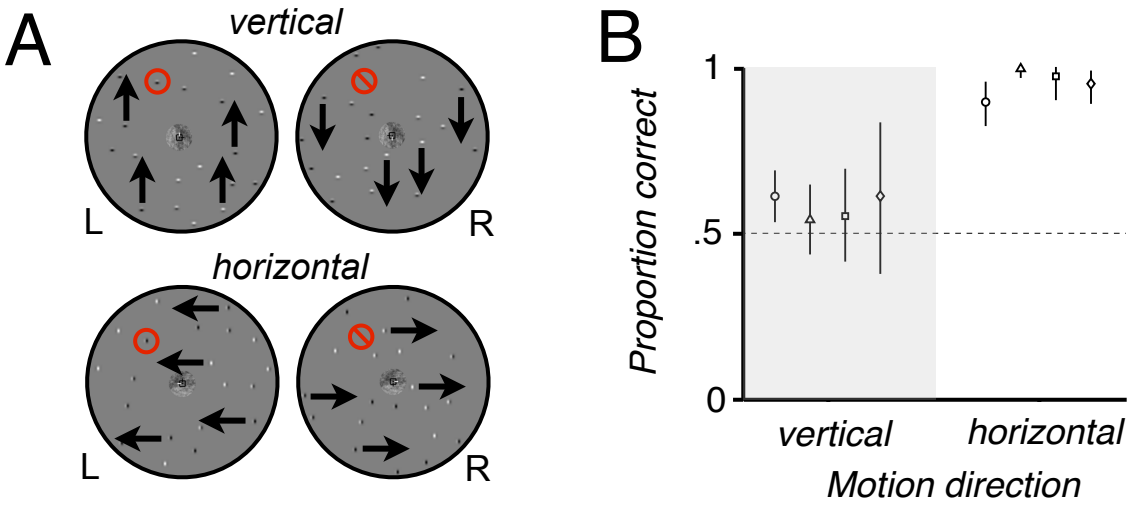

Figure A.4 Eye-of-origin & monocular motion control experiments

Control experiment to rule out possible joint eye-of-origin and monocular direction-of-motion discrimination. (A) Stimuli used to rule out joint eye-of-origin/direction-of-motion discrimination. All expert observers participated in a control experiment to address the concern that performance in the 3D motion direction discrimination task could have resulted from observers discriminating the direction of motion in one eye and mapping that to the correct 3D direction response (despite never receiving feedback). The 4 expert observers performed a 2AFC direction discrimination task on a dichoptic multiple-grating stimulus. This stimulus was similar to a dichoptic pseudoplaid, except all stimulus elements had the same orientation within each monocular view (as before, they moved in opposite directions in the two eyes). For the vertical motion (left panel), observers judged whether the motion in the right eye was upward or downward (as an assay of their ability to perform the posited joint eye-of-origin/direction-of-motion task); for the horizontal motion (right panel), they judged whether the 3D motion was "toward" or "away" (as an assay of their ability to discriminate 3D motion direction). (B) Proportion correct for vertical and horizontal motion discrimination tasks. As described in (A), motion of all stimulus elements was either vertical (left group of 4 points; each observer is a separate plotting symbol) or horizontal (right group). Symbols plot the average proportion correct from 90 trials per direction (error bars showing 95% confidence intervals). For all observers, accuracy was nearly perfect for the horizontal (3D direction) condition but close to chance for the vertical (joint eye-of-origin/direction-of-motion) condition. Observers could have used the component and/or pattern motion in this stimulus to perform the joint eye-of-origin/direction-of-motion task, but the data show little support for the use of either. These results are consistent with the modulation in performance as a function of direction in the main experiments. Most importantly, the far higher levels of 3D direction discrimination accuracy show that any sort of alternative strategy based on 2D (monocular) direction discrimination supported by simultaneous eye-of-origin discrimination is unlikely to account for our main findings.

Generalization of effects to naive observers and to a different task

We also confirmed that our key results (i.e., a dependence of 3D motion judgments parsimoniously explained by interocular comparisons of eye-specific pattern motions) could also be obtained from non-expert, naive subjects and in a simpler two-alternative forced-choice (2AFC) task. We therefore repeated the full set of dichoptic conditions (conventional gratings, Type I plaids, and Type II plaids, as well as both the two-component and multiple-component pseudoplaids) in 3 naive observers (one was an experienced psychophysical observer; the other two had no prior experience doing any sort of visual psychophysics) as they performed a 2AFC 3D motion direction discrimination task (indicating simply whether the stimulus appeared to move toward or away).

We observed a similar pattern of results in these 2AFC data collected in naive observers. The results of the conventional grating calibration experiment are shown in Figure A.5A. All observers showed a sinusoidal modulation of direction discrimination accuracy as a function of monocular direction, with a peak near horizontal (95% confidence interval, [–2, 3] deg). Figure A.5B is an ROC plot that relates these data to those from the corresponding rating scale task (refer back to Figure A.1B). This plot shows that sensitivity in the 2AFC task (individual square points) was highest for horizontal grating motions (darkest squares, $d' \sim 2$) and fell off for increasingly vertical grating motions (lighter squares, approaching d' of zero). The rating scale data are represented by connected curves and follow the same dependence on grating direction (higher for horizontal, lower for vertical), although the overall range of sensitivity spanned is wider and extends far above threshold (horizontal $d' \sim 5$).

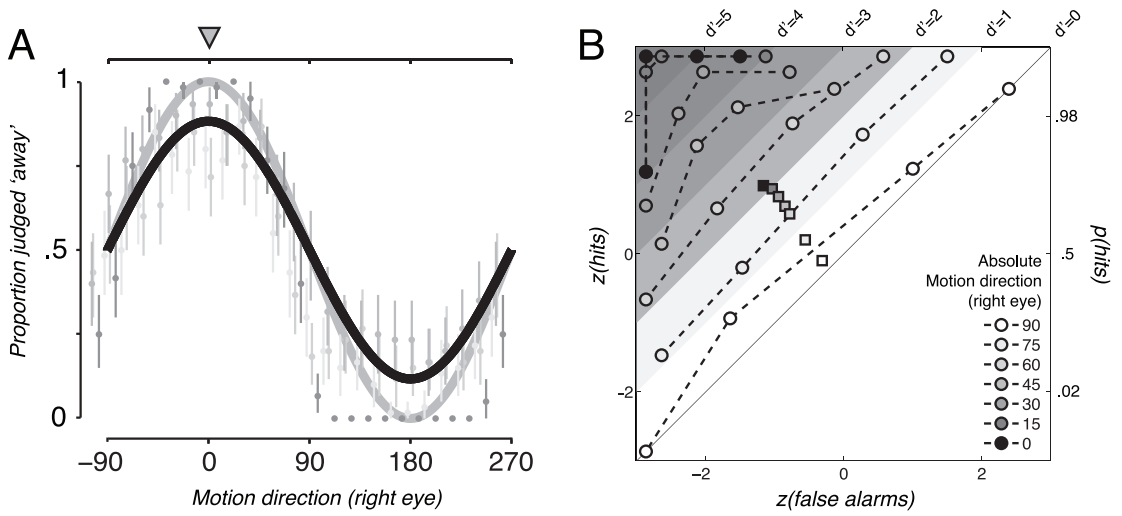

Figure A.5 Naive observer dichoptic grating results & ROC analysis

Naive observer results from a 2AFC task when viewing dichoptic gratings and their relationship to the experienced observer rating scale results in ROC space. (A) Three-dimensional direction discrimination accuracy in a 2AFC task, resulting from viewing a grating drifting in opposite direction in the two eyes for 3 naive and 1 expert observers. X-axis shows motion direction in the right eye (as in Figure A.1). Y-axis shows proportion of trials judged as moving "away" (the other trials were thus judged as "toward"); perfect performance would correspond to all trials between –90 and 90 judged as "away", and no trials between 90 and 270 judged as "away" (i.e., all trials judged as "toward"). Gray points show individual observer data. Error bars indicate the 95% confidence interval (data point darkness reflects amount of previous psychophysical experience). Darkest points indicating the one expert/author observer, and lightest points indicating an observer with no previous psychophysical experience. Black solid line reflects best cosine fit to the data from the three naive observers, i.e., not including the data from the expert/author observer. For reference, the gray solid line depicts the unit cosine function. The symbols above the upper x-axis indicate right eye motion direction for which the proportion of 3D "away" judgments would be largest if dependent on the horizontal aspect of the component motion (0 deg; gray solid triangle), and the best fit peak from the data (0 deg, 95% confidence interval, [–2, 3] deg; open black triangle—in this case, superimposed on gray triangle), replicating the pattern of results reported for the experienced observers. (B) Single grating ROC curves for the 4 expert observers (derived from their rating scale data) and 2AFC data from 3 naive observers. The abscissas and ordinates give the z-scores corresponding to the proportion of false alarms and hits, respectively (converting to z-scores linearizes the ROC curves assuming equal variance Gaussian noise on the decision axis). The use of a rating scale in expert observers allowed us to obtain reliable responses across a wide range of sensitivities (d′ ranges from 0 to 95), while avoiding ceiling and floor effects. Contours of equal d′ are given by the grayscale transitions. Circles connected by dashed lines show the averages

across the expert observers calculated from the rating scale data, with darker symbols denoting stimulus motion closer to horizontal ("Absolute motion direction" in legend indicates absolute angular deviation of monocular motion direction from horizontal). Estimated d′ and area under the curve both change systematically with motion direction. Squares show the corresponding 2AFC data for the 3 naive observers in the same ROC space. All three were close to chance for vertical motion and achieved a d′ greater than 2 for pure horizontal motion, with monotonically increasing performance in between.

The results from the other (plaid and pseudoplaid) experiments are shown in Figure A.6. Across all conditions, 3D motion direction discrimination sensitivity was highest when the pattern motion in each eye was horizontal (just as in the rating scale data). This was true for both the conventional Type I and Type II plaids (Figures A.6A and A.6B, respectively; 95% confidence intervals on the peak were [–3, 4] and [12, 20] deg) as well as for the two-component and multiple-component pseudoplaids (Figures A. 6C and A.6D; 95% confidence intervals were [–1, 8] and [–7, 7] deg). All of these 2AFC data can be explained by an interocular velocity difference mechanism that operates upon pattern motions for each eye, just as in the rating scale experiments.

The only quantitative difference between the original rating scale results and the these results is that the 2AFC confidence interval for Type II plaids ([12, 20] deg) excludes the predicted peak based on mean component direction (0 deg) or vector average (5.1 deg) but does not exclude one of the component motion vectors (15 deg). The rating scale confidence interval for expert observers did not contain either component motion vector. As for expert observers, the subsequent results for the naive observers with pseudoplaid stimuli can be used to establish the contribution of eye-specific pattern motion signals. Thus, these results overall further generalize the applicability of eye-specific pattern motions to a wider range of 3D motion phenomena and might support the relative utility of using rating scale tasks rooted in signal detection theory.

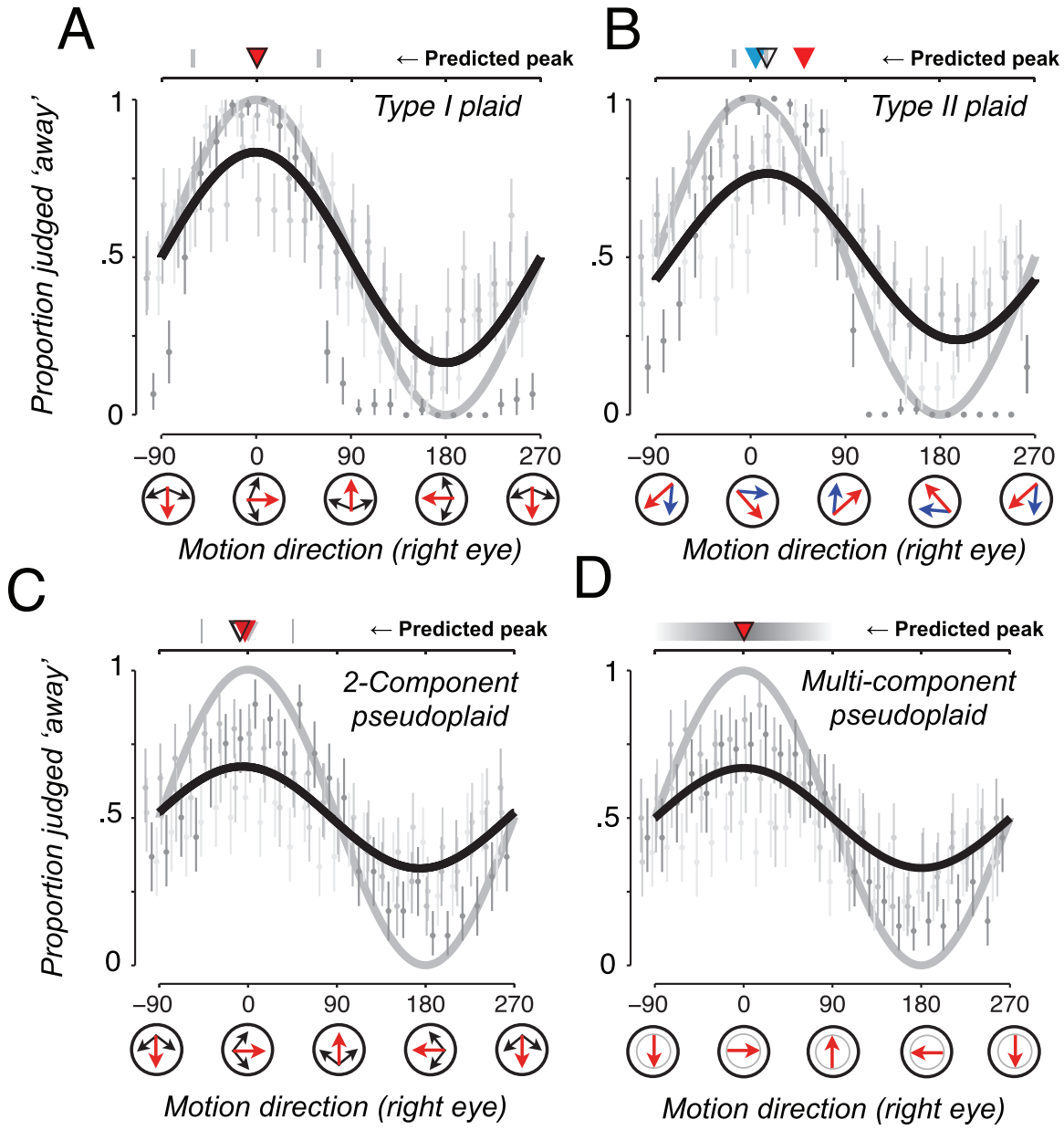

Figure A.6 Naive observer 2AFC results for plaids and pseudoplaids

(A) Three-dimensional direction discrimination accuracy resulting from viewing oppositely moving Type I plaids in the two eyes for 3 naive (and 1 expert) observers. Formatting as for Figures A.5A, A.2, and A.3. Symbols above the upper x-axis indicate predicted location of peak discrimination accuracy, based on either component motion (±60 deg; gray ticks) or pattern motion (0 deg; red solid triangle). Best fit to the data (0.1 deg, 95% confidence interval, [–3, 4] deg; open black triangle) indicates a dependence of accuracy on the plaid pattern motion direction. (B) Three-dimensional direction discrimination accuracy resulting from viewing oppositely moving Type II plaids in the two eyes for 3 naive (and 1 expert) observers. Formatting as above. Symbols above the

upper x-axis indicate predicted location of peak discrimination accuracy, based on either component motion (vector average 5.1 deg; blue solid triangle) or pattern motion (intersection of constraints (IOC) 51.2 deg; red solid triangle). The fitted peak of the perceptual tuning curve fell at 16.1 deg (black open triangle above plot; bounding 95% confidence interval, [12, 20] deg). (C) Three-dimensional direction discrimination accuracy resulting from viewing the 2-component pseudoplaid stimulus for 3 naive (and 1 expert) observers. Formatting as above. Three-dimensional motion direction discrimination accuracy depended on global pattern motion direction, even when individual stimulus elements were spaced so that the classical receptive field of a V1 only received input from one of the eyes. Peak accuracy at 1.1 deg (95% confidence interval on fitted peak location, [–1, 8] deg). (D) Three-dimensional direction discrimination accuracy resulting from viewing the multi-component pseudoplaid stimulus for 3 naive (and 1 expert) observers. Formatting as above. Three-dimensional motion direction discrimination accuracy depended on global pattern direction, even when individual stimulus elements were randomly oriented, with the constraint that each individual element's motion was compatible with a single global pattern motion direction (otherwise identical to the 2-component pseudoplaid). Peak at 0.3 deg (95% confidence interval on fitted peak location, [–7, 7] deg).

One final anecdote was also noted by the experimenters. Although all the expert/author observers reported rich 3D motion phenomenology when viewing the dichoptic pseudoplaids (as well as the other stimuli), the naive observers expressed varying strengths of subjective percepts—they initially claimed not to be able to see 3D motion in the more complex conditions (such as the pseudoplaids). One naive observer quickly came to report subjective experiences of 3D motion, one admitted to gaining some sort of an appreciation over time, and the other remained dubious to the end. Yet behavioral data from each of the 3 naive observers clearly demonstrate that they were able to discriminate 3D motion direction with good sensitivity (peaking at about $d' = 2$), despite an apparent dissociation from their phenomenology. Although this issue deserves more rigorous investigation, we note at this point that the somewhat mercurial perceptual experiences generated by the dichoptic pseudoplaids may hint at mechanisms that do not necessarily support strong perceptual experiences (at least without practice) but may still guide action effectively. Furthermore, rich subjective experiences of 3D motion from animated demonstrations of these stimuli may require patience from some observers. There may also be significant individual differences in the degree of reliance on the IOVD cue (Nefs et al., 2010). Figure A.7 contains a video demonstration of the pseudoplaid stimulus (along with an interpretive audio track).

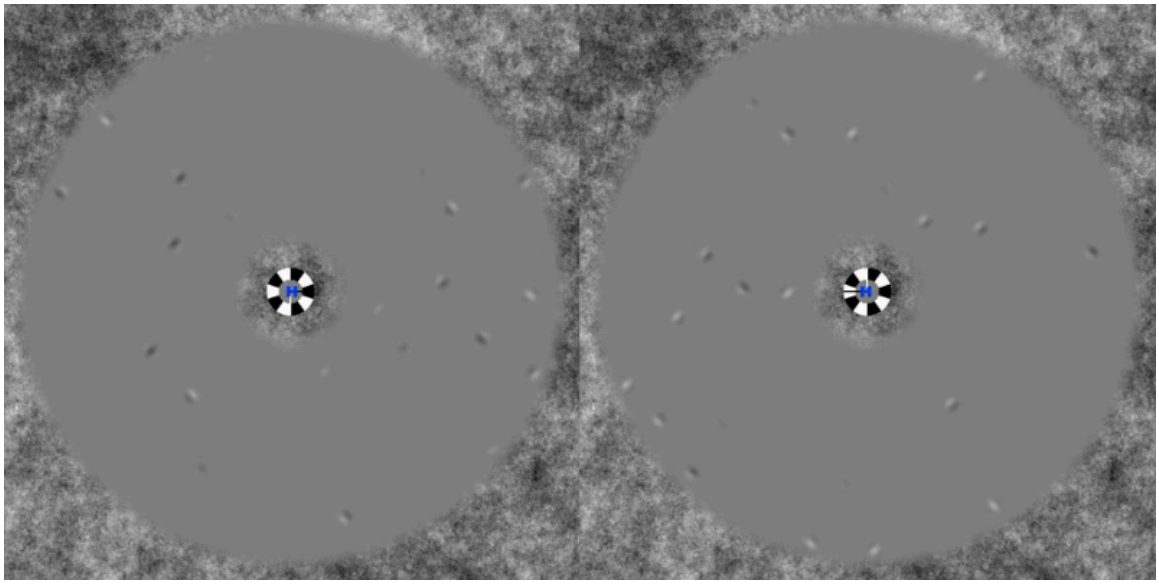

Figure A.7 Demonstration of the pseudoplaid stimulus (Movie 7)

Through either cross- or free-fusing one half of the image can be presented to one eye, and the other half presented to the other eye. Use your hands to block the undesired other half of the image in each eye for a cleaner percept. The movie presents a sequence of conditions building up to the novel pseudoplaid stimulus, which produces a percept of motion toward or away from the observer in the absence of binocular matches and thus in the absence of binocular disparity. Refer to the audio track within the movie for more details.

DISCUSSION

Our results motivate serious consideration of the proposition that 3D motion is computed by extracting the velocity difference between the pattern motion seen by each eye. Such a mechanism has some intriguing corollaries. First, interocular pattern motion differences per se support percepts of 3D motion: changing retinal disparities, as traditionally conceived, may be sufficient to yield percepts of 3D motion but are not necessary. Second, the binocular comparison (or "matching") that underlies the computation of these pattern-based interocular velocity differences occurs at a spatial scale much larger than that of a classical V1 receptive field. Finally, these large-scale pattern motions must also be "monocular" in the sense that a binocular comparison of them must be appropriately signed for the perception of "toward" versus "away" motion through depth.

Whenever observers dichoptically view complex spatiotemporal patterns, it is necessary to consider multiple potential sources of binocular and monocular information in interpreting the resulting perceptual sensitivities. The constellation of these results, most strongly demonstrated using the pseudoplaid stimuli, rules out explanations based on mechanisms other than pattern-motion-based IOVDs. First, one might wonder whether our results could be explained simply by sensitivity to changes in horizontal disparity. It is of course true that a horizontal disparity-based mechanism could explain the cosine dependence of discrimination performance in our preliminary grating experiment and also, perhaps, in the Type I experiment. However, such a mechanism would not explain the fact that Type II plaid sensitivity peaked outside the range of directions spanned by the individual components. Furthermore, conventional binocular disparities are extracted on the scale of V1 receptive fields and thus could not be used to account for the results of our dichoptic pseudoplaid experiments.

Second, psychophysical and physiological experiments have revealed sensitivity to disparities that are larger than those processed in V1 (Schor, Edwards, & Pope, 1998; Takemura, Inoue, Kawano, Quaia, & Miles, 2001). However, sensitivities to these large disparities have been shown to primarily reflect matching of temporally abrupt contrast

envelopes, with little or no contributions of the orientation-specific information within the envelopes (Wilcox & Allison, 2009). In contrast, our pseudoplaid Gabor elements were presented with gradual temporal onsets and within stationary envelopes. Because all 3D motion information was conveyed by the orientation and direction content within these envelopes, and not by the envelopes themselves, it is unlikely that even unconventionally large disparity mechanisms could explain our dichoptic pseudoplaid results. In particular, there is no way to explain the dependence on the direction of the pattern motion by appealing to disparities among the envelopes. The pseudoplaid data could be explained by suggesting a mechanism that compares the horizontal components of the velocities in each eye, but such a mechanism would be incompatible with the data from the preceding (conventional) Type II plaid experiment.

Finally, there is some evidence suggesting that binocular disparities can be extracted by comparing 2D spatial (mis)matches between the two eyes, instead of simply extracting the pure 1D horizontal component. Although such a mechanism could support the perceptual sensitivity we observed in our first set of grating and conventional plaid experiments, there were no systematic 2D matches between the left and right eyes' views in the dichoptic pseudoplaid experiments. In short, the dichoptic pseudoplaid stimulus provides such powerful evidence for an IOVD computation based on pattern motions because no known or posited disparity mechanism appears capable of extracting anything systematic from it. Furthermore, observers did not receive feedback, mitigating concerns that they may have arbitrarily mapped a particular eye's direction to a particular 3D motion response (a control experiment further weighs against this possibility, Figure A. 5B). Finally, we ensured that the starting, ending, and average disparities in our grating and plaid experiments were uninformative with respect to 3D direction, mitigating concerns that observers could have used a trivial static disparity cue.

One brief prior study investigated 3D motion percepts when viewing dichoptic gratings and Type I plaids (Wright & Gurney, 1992). Consistent with the results in our grating and conventional plaid experiments, they reported that 3D speed matches for dichoptic gratings followed the horizontal component of the velocity, and that stable

motion-in-depth percepts were generated by dichoptic Type I plaids, even when component grating orientations were close to vertical. They interpret their results as consistent with dependence on the pattern motion, and we agree with this interpretation. They also reported increasing perceived speed with increasingly near-vertical motion, while we report decreasing 3D motion discrimination performance as monocular motions approached vertical. These results are not inconsistent. Although we have not directly measured motion through depth discrimination as a function of pattern speed, we note that robust percepts of 3D motion occur at relatively slow pattern velocities—consistent with the binocular viewing geometry (wherein very fast environmental speeds in depth actually give rise to fairly slow projected speeds on the two retinae).

The notion of eye-specific pattern motion processing within large receptive fields has both computational and ecological appeals. Computationally, our results demonstrate that 3D motion signals are built directly from 2D pattern motion signals (and not from a larger number of ambiguous 1D signals). Ecologically, our results demonstrate that the visual system can use a global motion signal for each eye to compute the 3D direction of an object or surface, thus bypassing the traditional—and traditionally difficult—binocular matching problem. Thus, the visual system can still compute 3D direction even when different parts of an object or surface may be occluded for each eye, possibly at the expense of some spatial resolution.

These psychophysical results raise many questions and possibilities concerning how interocular pattern motions are computed in visual circuitry. One possibility is that an extrastriate area such as MT explicitly computes eyespecific pattern motions and/or their interocular difference (Zeki, 1974). However, individual neurons in MT lose their pattern motion sensitivity when 2-component pseudoplaids are placed within their receptive fields (Majaj et al., 2007), as well as when the individual component gratings making up a plaid are presented to separate eyes (Tailby et al., 2010). However, these stimuli differ in several regards (Gabor sizes, densities, speeds, etc.). Known physiology thus does not conclusively tell us whether MT would exhibit pattern motion selectivity to our dichoptic pseudoplaids.

Interocular comparison of pattern motion signals could be implemented explicitly in a pair of monocular pathways running through extrastriate dorsal cortex. Of course, this seems problematic given reports of only modest ocular dominance in extrastriate motion-sensitive cortical areas, such as MT (Deangelis & Newsome, 1999; Maunsell & Van Essen, 1983). On the other hand, the interocular comparison of pattern motion could still exploit these small degrees of ocularity (Sabatini, Solari, Andreani, Bartolozzi, & Bisio, 2001). This explanation assumes that, at the population level, small monocular biases are not ignored as "wiring noise" but are instead used to recover eye-of-origin information. Alternatively, interocular comparisons of pattern motions could be performed in the same step as the integration of multiple 1D signals; this possibility could be implemented without the existence of monocular neurons that are pattern motion selective.

We should note that our psychophysical stimuli (especially the dichoptic pseudoplaids) were different in key regards from the stimuli typically used to probe MT in electrophysiological experiments. For example, the vast majority of what we know about the function of area MT (such as pattern motion integration and binocularity) has been assessed using relatively fast frontoparallel speeds within the spatial receptive field of individual neurons (e.g. Maunsell & Van Essen, 1983). It remains logically possible that interocular pattern motion signals may only be evident when these neurons are tested using slower monocular speeds (like in our psychophysical stimuli), which are more consistent with the retinal projections of many ecologically valid 3D motions (Czuba et al., 2010). Likewise, our psychophysical stimuli has a larger spatial extent than a typical MT receptive field, and thus may have engaged surround mechanisms that are not as well understood (Raiguel, Van Hulle, Xiao, Marcar, & Orban, 1995; Xiao, Raiguel, Marcar, & Orban, 1997; 1998).

Alternatively, despite the fact that MT is considered a key stage in the computation of 2D pattern motion, other brain regions may be required to perform the motion integration demonstrated in our psychophysical results. Prior work has suggested degrees of pattern motion selectivity in V1 neurons (Guo, Benson, & Blakemore, 2004;

Tinsley et al., 2003), although debate surrounds details of experimental assessments of pattern motion selectivity (Movshon, Albright, Stoner, Majaj, & Smith, 2003; Pack, Berezovskii, & Born, 2001). Assuming this selectivity does exist, it is possible that pattern-based interocular velocity differences are computed in V1, although this still leaves open the question of how these pattern motions could be extracted at spatial scales larger than the classical receptive fields of V1 neurons. This alternate explanation must remain speculative until more is revealed about spatiotemporal integration in V1 neurons.

Another possibility is that eye-specific pattern motions might be extracted in a subcortical "blindsight" pathway that sends signals directly to dorsal extrastriate regions, including area MT (Barbur, Watson, Frackowiak, & Zeki, 1993; Berman & Wurtz, 2010; ffytche, Guy, & Zeki, 1995; Standage & Benevento, 1983). Motion perception preferentially survives in blindsight, and a tectofugal pathway might thus subserve some aspect of real-world (3D) motion perception.

Although it is difficult to psychophysically pinpoint the neural location of a computation with absolute certainty, our results do reveal that the nervous system somehow accomplishes a 3D motion computation that cannot be easily explained by the use of either early, monocular motion signals or any known or posited disparity mechanisms. The results motivate further study and use of the dichoptic pseudoplaid stimulus in both psychophysics and physiology. For example, one pressing question is whether the pattern motion computations involved in processing pseudoplaids are similar to those for conventional plaids. This is a topic of current study in our laboratory, involving the comparison of Type I and Type II pseudoplaids, as well as the consideration of spatial and temporal parameters that may affect the pattern motion computations (Takeuchi, 1998).

More broadly, these results suggest a fresh perspective for thinking about a range of previous findings. Several studies have demonstrated a role for interocular velocity differences in 3D motion perception (Beverley & Regan, 1973; Brooks, 2002b; Fernandez & Farell, 2006; J. M. Harris & Rushton, 2003; Rokers et al., 2009; Shioiri et al., 2000), as well as contributions of monocularly occluded stimuli (Brooks & Gillam,

2006; 2007), but these studies did not directly investigate the level of the motion processing hierarchy at which this computation occurs. (One recent study did report that the IOVD mechanism operates on eye-specific motion signals that are broadband in spatial frequency, suggestive of a later stage of spatiotemporal integration (Shioiri et al., 2009).) Meanwhile, a separate line of work has suggested monocular contributions to 2D pattern motion perception (Alais, Burke, & Wenderoth, 1996a; Alais, van der Smagt, Verstraten, & van de Grind, 1996b; Burke & Wenderoth, 1993) but did not consider the potential functional role of these signals in 3D motion perception. It may be possible to parsimoniously integrate these seemingly disparate lines of work with a single appeal to eye-specific pattern motions. Given that the brain appears to compute and compare eye-specific pattern motions for recovering 3D motion, the challenge now is to understand how and where such a computation occurs (Likova & Tyler, 2007; Rokers et al., 2009).

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, Optics, image science, and vision*, *2*(2), 284–299.

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525.

Akase, E., Inokawa, H., & Toyama, K. (1998). Neuronal responsiveness to three-dimensional motion in cat posteromedial lateral suprasylvian cortex. *Experimental brain research*, *122*(2), 214–226.

Alais, D., Burke, D., & Wenderoth, P. (1996a). Further evidence for monocular determinants of perceived plaid direction. *Vision Research*, *36*(9), 1247–1253.

Alais, D., van der Smagt, M. J., Verstraten, F. A., & van de Grind, W. A. (1996b). Monocular mechanisms determine plaid motion coherence. *Visual Neuroscience*, *13*(4), 615–626.

Alais, D., Verstraten, F. A. J., & Burr, D. C. (2005). The motion aftereffect of transparent motion: two temporal channels account for perceived direction. *Vision Research*, *45*(4), 403–412. doi:10.1016/j.visres.2004.09.005

Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *Journal of Vision*, *9*(3), 4.1–25. doi:10.1167/9.3.4

Andrews, T. J., Glennerster, A., & Parker, A. J. (2001). Stereoacuity thresholds in the presence of a reference surface. *Vision Research*, *41*(23), 3051–3061.

Anstis, S., & Duncan, K. (1983). Separate motion aftereffects from each eye and from both eyes. *Vision Research*, *23*(2), 161–169.

Anstis, S., Giaschi, D., & Cogan, A. I. (1985). Adaptation to apparent motion. *Vision Research*, *25*(8), 1051–1062.

Anstis, S., Verstraten, F. A., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences*, *2*(3), 111–117.

Barbur, J. L., Watson, J. D., Frackowiak, R. S., & Zeki, S. (1993). Conscious visual perception without V1. *Brain*, *116*(6), 1293–1302.

Barlow, H. B., & HILL, R. M. (1963). Evidence for a Physiological Explanation of the Waterfall Phenomenon and Figural After-effects. *Nature*, *200*, 1345–1347.

Berman, R. A., & Wurtz, R. H. (2010). Functional identification of a pulvinar path from superior colliculus to cortical area MT. *The Journal of Neuroscience*, *30*(18), 6342–6354. doi:10.1523/JNEUROSCI.6176-09.2010

Beverley, K. I., & Regan, D. (1973). Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *The Journal of Physiology*, *235*(1), 17–29.

Blake, R., & Hiris, E. (1993). Another means for measuring the motion aftereffect. *Vision Research*, *33*(11), 1589–1592.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *The Journal of Physiology*, *211*(3), 599–622.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, *203*(1), 237–260.

Born, R. T., & bradley, D. C. (2005). Structure and function of visual area MT. *Annual review of neuroscience*, *28*, 157–189. doi:10.1146/annurev.neuro.26.041002.131052

Bowd, C., Donnelly, M., Shorter, S., & Patterson, R. (2000). Cross-domain adaptation reveals that a common mechanism computes stereoscopic (cyclopean) and luminance plaid motion. *Vision Research*, *40*(3), 331–339.

Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research*, *14*(7), 519–527.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience*, *12*(12), 4745–4765.

Brooks, K. R. (2002a). Monocular motion adaptation affects the perceived trajectory of stereomotion. *Journal of experimental psychology Human perception and performance*, *28*(6), 1470–1482.

Brooks, K. R. (2002b). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision*, *2*(3), 218–231. doi:10:1167/2.3.2

Brooks, K. R., & Gillam, B. J. (2006). The swinging doors of perception: stereomotion without binocular matching. *Journal of Vision*, *6*(7), 685–695. doi:10.1167/6.7.2

Brooks, K. R., & Gillam, B. J. (2007). Stereomotion perception for a monocularly camouflaged stimulus. *Journal of Vision*, *7*(13), 1.1–14. doi:10.1167/7.13.1

Brooks, K. R., & Mather, G. (2000). Perceived speed of motion in depth is reduced in the periphery. *Vision Research*, *40*(25), 3507–3516.

Brooks, K. R., & Stone, L. S. (2004). Stereomotion speed perception: contributions from both changing disparity and interocular velocity difference over a range of relative disparities. *Journal of Vision*, *4*(12), 1061–1079. doi:10:1167/4.12.6

Brooks, K. R., & Stone, L. S. (2006a). Stereomotion suppression and the perception of speed: accuracy and precision as a function of 3D trajectory. *Journal of Vision*, *6* (11), 1214–1223. doi:10.1167/6.11.6

Brooks, K. R., & Stone, L. S. (2006b). Spatial scale of stereomotion speed processing. *Journal of Vision*, *6*(11), 1257–1266. doi:10.1167/6.11.9

Bunday, K. L., & Bronstein, A. M. (2008). Visuo-vestibular influences on the moving platform locomotor aftereffect. *Journal of Neurophysiology*, *99*(3), 1354–1365. doi:10.1152/jn.01214.2007

Burke, D., & Wenderoth, P. (1993). Determinants of two-dimensional motion aftereffects induced by simultaneously- and alternately-presented plaid components. *Vision Research*, *33*(3), 351–359.

Burr, D. C., & Santoro, L. (2001). Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Research*, *41*(15), 1891–1899.

Burr, D., & Thompson, P. (2011). Motion psychophysics: 1985-2010. *Vision Research*, *51* (13), 1431–1456. doi:10.1016/j.visres.2011.02.008

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science (New York, NY)*, *208*(4444), 615–617. doi:10.1126/science.7367885

Cameron, E. L., Baker, C. L., & Boulton, J. C. (1992). Spatial frequency selective mechanisms underlying the motion aftereffect. *Vision Research*, *32*(3), 561–568.

Carney, T., & Shadlen, M. N. (1992). Binocularity of early motion mechanisms: comments on Georgeson and Shackleton. *Vision Research*, *32*(1), 187–191.

Carney, T., & Shadlen, M. N. (1993). Dichoptic activation of the early motion system. *Vision Research*, *33*(14), 1977–1995.

Clark, A. M., & bradley, D. C. (2008). Integration of distributed one-dimensional motion signals by macaque middle temporal cortical neurons. *Computational & Systems Neuroscience (COSYNE) Conference*, 1–1.

Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., & Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli. *Perception*, *24*(1), 33–47.

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, *31*(12), 2195–2207.

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1993). Disparity-tuned channels of the human visual system. *Visual neuroscience*, *10*(4), 585–596.

Cumming, B. G. (1995). The relationship between stereoacuity and stereomotion thresholds. *Perception*, *24*(1), 105–114.

Cumming, B. G., & Parker, A. J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research*, *34*(4), 483–495.

Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, *27*(11), 1367–1377.

Cynader, M., & Regan, D. (1978). Neurones in cat parastriate cortex sensitive to the direction of motion in three-dimensional space. *The Journal of Physiology*, *274*, 549–569.

Cynader, M., & Regan, D. (1982). Neurons in cat visual cortex tuned to the direction of motion in depth: effect of positional disparity. *Vision Research*, *22*(8), 967–982.

Czuba, T. B., Huk, A. C., & Cormack, L. K. (2011). Isolation of binocular 3D motion cues in human visual cortex. *Journal of Vision*, *11*(11), 710–710. doi: 10.1167/11.11.710

Czuba, T. B., Rokers, B., Huk, A. C., & Cormack, L. K. (2012). To CD or not to CD: Is there a 3D motion aftereffect based on changing disparities? *Journal of Vision*, *12* (4). doi:10.1167/12.4.7

Czuba, T. B., Rokers, B., Guillet, K., Huk, A. C., & Cormack, L. K. (2011). Three-dimensional motion aftereffects reveal distinct direction-selective mechanisms for binocular processing of motion through depth. *Journal of Vision*, *11*(10), 18. doi: 10.1167/11.10.18

Czuba, T. B., Rokers, B., Huk, A. C., & Cormack, L. K. (2010). Speed and eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion. *Journal of Neurophysiology*, *104*(5), 2886–2899. doi:10.1152/jn.00585.2009

Deangelis, G. C., & Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *The Journal of Neuroscience*, *19*(4), 1398–1415.

Deangelis, G. C., & Newsome, W. T. (2004). Perceptual "read-out" of conjoined direction and disparity maps in extrastriate area MT. *PLoS biology*, *2*(3), E77. doi:10.1371/journal.pbio.0020077

Delicato, L. S., & Qian, N. (2005). Is depth perception of stereo plaids predicted by intersection of constraints, vector average or second-order feature? *Vision Research*, *45*(1), 75–89. doi:10.1016/j.visres.2004.07.028

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, *3*(10), 586–598. doi:10:1167/3.10.1

Duhamel, J. R., Bremmer, F., BenHamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, *389*(6653), 845–848. doi:10.1038/39865

Durgin, F. H., & Hammer, J. T. (2001). Visual aftereffects of sequential perception: dynamic adaptation to changes in texture density and contrast. *Vision Research*, *41*(20), 2607–2617.

Farell, B. (2003). Detecting disparity in two-dimensional patterns. *Vision Research*, *43* (9), 1009–1026.

Farell, B., Chai, Y.-C., & Fernandez, J. M. (2009). Projected disparity, not horizontal disparity, predicts stereo depth of 1-D patterns. *Vision Research*, *49*(17), 2209–2216. doi:10.1016/j.visres.2009.06.013

Fernandez, J. M., & Farell, B. (2005). Seeing motion in depth using inter-ocular velocity differences. *Vision Research*, *45*(21), 2786–2798. doi:10.1016/j.visres.2005.05.021

Fernandez, J. M., & Farell, B. (2006). Motion in depth from interocular velocity differences revealed by differential motion aftereffect. *Vision Research*, *46*(8-9), 1307–1317. doi:10.1016/j.visres.2005.10.025

ffytche, D. H., Guy, C. N., & Zeki, S. (1995). The parallel visual motion inputs into areas V1 and V5 of human cerebral cortex. *Brain : a journal of neurology*, *118 (Pt 6)*, 1375–1394.

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *The Journal of Neuroscience*, *28*(15), 3988–3999. doi:10.1523/JNEUROSCI.5476-07.2008

Georgeson, M. A., & Schofield, A. J. (2002). Shading and texture: separate information channels with a common adaptation mechanism? *Spatial Vision*, *16*(1), 59–76.

Gibson, J. J. (1950). *The perception of the visual world* (p. 235).

Gray, R., & Regan, D. (1996). Cyclopean motion perception produced by oscillations of size, disparity and location. *Vision Research*, *36*(5), 655–665.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. John Wiley & Sons.

Grunewald, A., & Lankheet, M. J. M. (1996). Orthogonal motion after-effect illusion predicted by a model of cortical motion processing. *Nature*, *384*(6607), 358–360. doi:10.1038/384358a0

Grunewald, A., & Mingolla, E. (1998). Motion after-effect due to binocular sum of adaptation to linear motion. *Vision Research*, *38*(19), 2963–2971.

Guo, K., Benson, P. J., & Blakemore, C. (2004). Pattern motion is present in V1 of awake but not anaesthetized monkeys. *The European journal of neuroscience*, *19*(4), 1055–1066.

Harris, J. M., & Dean, P. J. A. (2003). Accuracy and precision of binocular 3-D motion perception. *Journal of experimental psychology Human perception and performance*, *29*(5), 869–881. doi:10.1037/0096-1523.29.5.869

Harris, J. M., & Rushton, S. K. (2003). Poor visibility of motion in depth is due to early motion averaging. *Vision Research*, *43*(4), 385–392.

Harris, J. M., & Watamaniuk, S. N. (1995). Speed discrimination of motion-in-depth using binocular cues. *Vision Research*, *35*(7), 885–896.

Harris, J. M., Nefs, H. T., & Grafton, C. E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, *21*(6), 531–547. doi:10.1163/156856808786451462

Harris, L. R., Morgan, M. J., & Still, A. W. (1981). Moving and the motion after-effect. *Nature*, *293*(5828), 139–141.

Hess, R. F., Hutchinson, C. V., Ledgeway, T., & Mansouri, B. (2007). Binocular influences on global motion processing in the human visual system. *Vision Research*, *47*(12), 1682–1692. doi:10.1016/j.visres.2007.02.005

Hiris, E., & Blake, R. (1992). Another perspective on the visual motion aftereffect. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(19), 9025–9028.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *The Journal of Neuroscience*, *22*(16), 7195–7205. doi:20026661

Huk, A. C., Ress, D., & Heeger, D. J. (2001). Neuronal basis of the motion aftereffect reconsidered. *Neuron*, *32*(1), 161–172.

Jazayeri, M., & Movshon, J. A. (2007). Integration of sensory evidence in motion discrimination. *Journal of Vision*, *7*(12), 7.1–7. doi:10.1167/7.12.7

Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell System Technical Journal*.

Julesz, B. (1971). Foundations of cyclopean perception.

Julesz, B., & Payne, R. A. (1968). Differences between monocular and binocular stroboscopic movement perception. *Vision Research*, *8*(4), 433–444.

Khawaja, F. A., Tsui, J. M. G., & Pack, C. C. (2009). Pattern motion selectivity of spiking outputs and local field potentials in macaque visual cortex. *The Journal of Neuroscience*, *29*(43), 13702–13709. doi:10.1523/JNEUROSCI.2844-09.2009

Knapen, T., Rolfs, M., & Cavanagh, P. (2009). The reference frame of the motion aftereffect is retinotopic. *Journal of Vision*, *9*(5), 16.1–7. doi:10.1167/9.5.16

Kohn, A., & Movshon, J. A. (2003). Neuronal adaptation to visual motion in area MT of the macaque. *Neuron*, *39*(4), 681–691.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Publishing Group*, *3*(12), 1340–1345. doi:10.1038/81887

Lankheet, M. J. M., van Doorn, A. J., Bouman, M. A., & van de Grind, W. A. (2000). Motion coherence detection as a function of luminance level in human central vision. *Vision Research*, *40*(26), 3599–3611.

Lankheet, M. J., & Palmen, M. (1998). Stereoscopic segregation of transparent surfaces and the effect of motion contrast. *Vision Research*, *38*(5), 659–668.

Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *The Journal of Neuroscience*, *26*(51), 13128–13142. doi:10.1523/JNEUROSCI.1657-06.2006

Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology*, *95*(2), 862–881. doi:10.1152/jn.00668.2005

Lee, H. A., & Lee, S.-H. (2012). Hierarchy of direction-tuned motion adaptation in human visual cortex. *Journal of Neurophysiology*, *107*(8), 2163–2184. doi:10.1152/jn.00923.2010

Lehmkuhle, S. W., & Fox, R. (1976). On measuring interocular transfer. *Vision Research*, *16*(4), 428–430.

Likova, L. T., & Tyler, C. W. (2007). Stereomotion processing in the human occipital cortex. *NeuroImage*, *38*(2), 293–305. doi:10.1016/j.neuroimage.2007.06.039

Liu, J., & Newsome, W. T. (2003). Functional organization of speed tuned neurons in visual area MT. *Journal of Neurophysiology*, *89*(1), 246–256. doi:10.1152/jn.00097.2002

Liu, L., Tyler, C. W., & Schor, C. M. (1992). Failure of rivalry at low contrast: evidence of a suprathreshold binocular summation process. *Vision Research*, *32*(8), 1471–1479.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (2nd ed. p. 512). Psychology Press.

Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, *2*, 49–55.

Majaj, N. J., Carandini, M., & Movshon, J. A. (2007). Motion integration by neurons in macaque MT is local, not global. *The Journal of Neuroscience*, *27*(2), 366–370. doi:10.1523/JNEUROSCI.3183-06.2007

Mather, G. (1980). The movement aftereffect and a distribution-shift model for coding the direction of visual movement. *Perception*, *9*(4), 379–392.

Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, *49*(5), 1148–1167.

McKee, S. P., & Nakayama, K. (1984). The detection of motion in the peripheral visual field. *Vision Research*, *24*(1), 25–32.

Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology*, *55*(6), 1308–1327.

Mitchell, D. E., Reardon, J., & Muir, D. W. (1975). Interocular transfer of the motion after-effect in normal and stereoblind observers. *Experimental brain research*, *22* (2), 163–173.

Mollon, J. D. (1974). After-effects and the brain. *New Scientist*, *61*, 479–482.

Morgan, M. J. (2012). Motion adaptation does not depend on attention to the adaptor. *Vision Research*, *55*, 47–51. doi:10.1016/j.visres.2011.12.009

Morgan, M. L., Deangelis, G. C., & Angelaki, D. E. (2008). Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron*, *59*(4), 662–673. doi: 10.1016/j.neuron.2008.06.024

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. (C. Chagas, R. Gattass, & C. Gross, Eds.)*Pattern recognition mechanisms*, *54*, 117–151.

Movshon, J. A., Albright, T. D., Stoner, G. R., Majaj, N. J., & Smith, M. A. (2003). Cortical responses to visual motion in alert and anesthetized monkeys. *Nature Publishing Group*, *6*(1), 3; author reply 3–4. doi:10.1038/nn0103-3a

Nefs, H. T., OHare, L., & Harris, J. M. (2010). Two independent mechanisms for motion-in-depth perception: evidence from individual differences. *Frontiers in Perception Science*, 1–8.

Neri, P., Bridge, H., & Heeger, D. J. (2004). Stereoscopic processing of absolute and relative disparity in human visual cortex. *Journal of Neurophysiology*, *92*(3), 1880–1891. doi:10.1152/jn.01042.2003

Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, *401*(6754), 695–698. doi:10.1038/44409

Nestares, O., & Heeger, D. J. (2000). Robust multiresolution alignment of MRI brain volumes. *Magnetic Resonance in Medicine*, *43*(5), 705–715.

Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, *8*(6), 2201–2211.

Norcia, A. M., & Tyler, C. W. (1984). Temporal frequency limits for stereoscopic apparent motion processes. *Vision Research*, *24*(5), 395–401.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science (New York, NY)*, *249*(4972), 1037–1041.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1996). Encoding of binocular disparity by simple cells in the cat's visual cortex. *Journal of Neurophysiology*, *75*(5), 1779–1805.

Ono, H., & Barbeito, R. (1985). Utrocular discrimination is not sufficient for utrocular identification. *Vision Research*, *25*(2), 289–299.

Orban, G. A., Kennedy, H., & Bullier, J. (1986). Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: influence of eccentricity. *Journal of Neurophysiology*, *56*(2), 462–480.

Pack, C. C., Berezovskii, V. K., & Born, R. T. (2001). Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature*, *414*(6866), 905–908. doi:10.1038/414905a

Palmisano, S., Allison, R. S., & Howard, I. P. (2006). Effect of decorrelation on 3-D grating detection with static and dynamic random-dot stereograms. *Vision Research*, *46*(1-2), 57–71. doi:10.1016/j.visres.2005.10.005

Patterson, R. (2009). Unresolved issues in stereopsis: dynamic disparity processing. *Spatial Vision*, *22*(1), 83–90. doi:10.1163/156856809786618510

Patterson, R., Bowd, C., Phinney, R., Fox, R., & Lehmkuhle, S. W. (1996). Disparity tuning of the stereoscopic (cyclopean) motion aftereffect. *Vision Research*, *36*(7), 975–983.

Patterson, R., Bowd, C., Phinney, R., Pohndorf, R., Barton-Howard, W. J., & Angilletta, M. (1994). Properties of the stereoscopic (cyclopean) motion aftereffect. *Vision Research*, *34*(9), 1139–1147.

Patterson, R., Fournier, L. R., Wiediger, M., Vavrek, G., Becker-Dippman, C., & Bickler, I. (2005). Selective attention and cyclopean motion processing. *Vision Research*, *45*(20), 2601–2607. doi:10.1016/j.visres.2005.03.003

Peng, Q., & Shi, B. E. (2010). The changing disparity energy model. *Vision Research*, *50* (2), 181–192. doi:10.1016/j.visres.2009.11.012

Perge, J. A., Borghuis, B. G., Bours, R. J. E., Lankheet, M. J. M., & van Wezel, R. J. A. (2005). Temporal dynamics of direction tuning in motion-sensitive macaque area MT. *Journal of Neurophysiology*, *93*(4), 2104–2116. doi:10.1152/jn.00601.2004

Perrone, J. A., & Thiele, A. (2002). A model of speed tuning in MT neurons. *Vision Research*, *42*(8), 1035–1051.

Pianta, M. J., & Gillam, B. J. (2003). Paired and unpaired features can be equally effective in human depth perception. *Vision Research*, *43*(1), 1–6.

Poggio, G. F., & Poggio, T. (1984). The analysis of stereopsis. *Annual review of neuroscience*, *7*, 379–412. doi:10.1146/annurev.ne.07.030184.002115

Poggio, G. F., & Talbot, W. H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *The Journal of Physiology*, *315*, 469–492.

Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *The Journal of Neuroscience*, *8*(12), 4531–4550.

Poggio, G. F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, *25*(3), 397–406.

Porac, C., & Coren, S. (1986). Sighting dominance and utrocular discrimination. *Perception and Psychophysics*, *39*(6), 449–451.

Portfors-Yeomans, C. V., & Regan, D. (1996). Cyclopean discrimination thresholds for the direction and speed of motion in depth. *Vision Research*, *36*(20), 3265–3279.

Portfors-Yeomans, C. V., & Regan, D. (1997). Discrimination of the direction and speed of motion in depth of a monocularly visible target from binocular information alone. *Journal of experimental psychology Human perception and performance*, *23*(1), 227–243.

Priebe, N. J., Lisberger, S. G., & Movshon, J. A. (2006). Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *The Journal of Neuroscience*, *26*(11), 2941–2950. doi:10.1523/JNEUROSCI.3936-05.2006

Prince, S. J. D., Cumming, B. G., & Parker, A. J. (2002). Range and mechanism of encoding of horizontal disparity in macaque V1. *Journal of Neurophysiology*, *87*(1), 209–221.

Raiguel, S., Van Hulle, M. M., Xiao, D. K., Marcar, V. L., & Orban, G. A. (1995). Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. *The European journal of neuroscience*, *7*(10), 2064–2082.

Raymond, J. E. (1993). Complete interocular transfer of motion adaptation effects on motion coherence thresholds. *Vision Research*, *33*(13), 1865–1870.

Read, J. C. A., Phillipson, G. P., Serrano-Pedraza, I., Milner, A. D., & Parker, A. J. (2010). Stereoscopic vision in the absence of the lateral occipital cortex. *PloS one*, *5*(9), e12608. doi:10.1371/journal.pone.0012608

Regan, D., & Beverley, K. I. (1973). Some dynamic features of depth perception. *Vision Research*, *13*(12), 2369–2379.

Regan, D., & Cynader, M. (1982). Neurons in cat visual cortex tuned to the direction of motion in depth: effect of stimulus speed. *Investigative Ophthalmology & Visual Science*, *22*(4), 535–550.

Regan, D., & Gray, R. (2000). Visually guided collision avoidance and collision achievement. *Trends in Cognitive Sciences*, *4*(3), 99–107.

Regan, D., & Gray, R. (2009). Binocular processing of motion: some unresolved questions. *Spatial Vision*, *22*(1), 1–43. doi:10.1163/156856809786618501

Richards, W. (1972). Response functions for sine-and square-wave modulations of disparity. *Journal of the Optical Society of America*, *62*(7), 907–911.

Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental brain research*, *75*(1), 53–64.

Rokers, B., Cormack, L. K., & Huk, A. C. (2008a). Strong percepts of motion through depth without strong percepts of position in depth. *Journal of Vision*, *8*(4), 6.1–10. doi:10.1167/8.4.6

Rokers, B., Cormack, L. K., & Huk, A. C. (2008b). Neural circuits underlying the perception of 3D motion. *VSS 2008*, 1.

Rokers, B., Cormack, L. K., & Huk, A. C. (2009). Disparity- and velocity-based signals for three-dimensional motion perception in human MT. *Nature Neuroscience*, *12* (8), 1050–1055. doi:10.1038/nn.2343

Rokers, B., Czuba, T. B., Cormack, L. K., & Huk, A. C. (2011). Motion processing with two eyes in three dimensions. *Journal of Vision*, *11*(2). doi:10.1167/11.2.10

Rovamo, J., & Raninen, A. (1984). Critical flicker frequency and M-scaling of stimulus size and retinal illuminance. *Vision Research*, *24*(10), 1127–1131.

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Publishing Group*, *9*(11), 1421–1431. doi:10.1038/nn1786

Sabatini, S. P., Solari, F., Andreani, G., Bartolozzi, C., & Bisio, G. M. (2001). A hierarchical model of complex cells in visual cortex for the binocular perception of motion-in-depth. *Advances in Neural Information Processing Systems*, 1271–1278.

Sakano, Y., Allison, R. S., & Howard, I. P. (2005). Aftereffects of motion in depth based on binocular cues. *Journal of Vision*, *5*(8), 732–732. doi:10.1167/5.8.732

Sakano, Y., Allison, R. S., & Howard, I. P. (2012). Motion aftereffect in depth based on binocular information. *Journal of Vision*, *12*(1). doi:10.1167/12.1.11

Sakano, Y., Allison, R. S., Howard, I. P., & Sadr, S. (2006). Aftereffect of motion-in-depth based on binocular cues: No effect of relative disparity between adaptation and test surfaces. *Journal of Vision*, *6*(6), 626–626. doi:10.1167/6.6.626

Scase, M. O., Braddick, O. J., & Raymond, J. E. (1996). What is noise for the motion system? *Vision Research*, *36*(16), 2579–2586.

Schor, C. M., & Badcock, D. R. (1985). A comparison of stereo and vernier acuity within spatial channels as a function of distance from fixation. *Vision Research*, *25*(8), 1113–1119.

Schor, C. M., Edwards, M., & Pope, D. R. (1998). Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Research*, *38*(20), 3057–3068.

Shadlen, M. N., & Carney, T. (1986). Mechanisms of human motion perception revealed by a new cyclopean illusion. *Science (New York, NY)*, *232*(4746), 95–97.

Shioiri, S., & Matsumiya, K. (2009). Motion mechanisms with different spatiotemporal characteristics identified by an MAE technique with superimposed gratings. *Journal of Vision*, *9*(5), 30.1–15. doi:10.1167/9.5.30

Shioiri, S., Kakehi, D., Tashiro, T., & Yaguchi, H. (2009). Integration of monocular motion signals and the analysis of interocular velocity differences for the perception of motion-in-depth. *Journal of Vision*, *9*(13), 10.1–17. doi: 10.1167/9.13.10

Shioiri, S., Nakajima, T., Kakehi, D., & Yaguchi, H. (2008). Differences in temporal frequency tuning between the two binocular mechanisms for seeing motion in depth. *Journal of the Optical Society of America A, Optics, image science, and vision*, *25*(7), 1574–1585.

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, *40*(19), 2565–2572.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743–761.

Standage, G. P., & Benevento, L. A. (1983). The organization of connections between the pulvinar and visual area MT in the macaque monkey. *Brain research*, *262*(2), 288–294.

Stevenson, S. B., Cormack, L. K., & Schor, C. M. (1994). The effect of stimulus contrast and interocular correlation on disparity vergence. *Vision Research*, *34*(3), 383–396.

Stevenson, S. B., Cormack, L. K., Schor, C. M., & Tyler, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vision Research*, *32*(9), 1685–1694.

Tailby, C., Majaj, N. J., & Movshon, J. A. (2010). Binocular integration of pattern motion signals by MT neurons and by human observers. *The Journal of Neuroscience*, *30* (21), 7344–7349. doi:10.1523/JNEUROSCI.4552-09.2010

Takemura, H., Inoue, Y., Kawano, K., Quaia, C., & Miles, F. A. (2001). Single-unit activity in cortical area MST associated with disparity-vergence eye movements: evidence for population coding. *Journal of Neurophysiology*, *85*(5), 2245–2266.

Takeuchi, T. (1998). Effect of contrast on the perception of moving multiple Gabor patterns. *Vision Research*, *38*(20), 3069–3082.

Tao, R., Lankheet, M. J. M., van de Grind, W. A., & van Wezel, R. J. A. (2003). Velocity dependence of the interocular transfer of dynamic motion aftereffects. *Perception*, *32*(7), 855–866.

Tinsley, C. J., Webb, B. S., Barraclough, N. E., Vincent, C. J., Parker, A., & Derrington, A. M. (2003). The nature of V1 neural responses to 2D moving patterns depends on receptive-field structure in the marmoset monkey. *Journal of Neurophysiology*, *90*(2), 930–937. doi:10.1152/jn.00708.2002

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal of Neuroscience*, *15*(4), 3215–3230.

Toyama, K., Komatsu, Y., Kasai, H., Fujii, K., & Umetani, K. (1985). Responsiveness of Clare-Bishop neurons to visual cues associated with motion of a visual stimulus in three-dimensional space. *Vision Research*, *25*(3), 407–414.

Tyler, C. W. (1971). Stereoscopic depth movement: two eyes less sensitive than one. *Science (New York, NY)*, *174*(4012), 958–961.

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research*, *15*(5), 583–590.

Tyler, C. W., Baseler, H. A., Kontsevich, L. L., Likova, L. T., Wade, A. R., & Wandell, B. A. (2005). Predominantly extra-retinotopic cortical response to pattern symmetry. *NeuroImage*, *24*(2), 306–314. doi:10.1016/j.neuroimage.2004.09.018

van de Grind, W. A., Lankheet, M. J. M., & Tao, R. (2003). A gain-control model relating nulling results to the duration of dynamic motion aftereffects. *Vision Research*, *43*(2), 117–133.

Van Essen, D. C., Maunsell, J. H., & Bixby, J. L. (1981). The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *The Journal of comparative neurology*, *199*(3), 293–326. doi:10.1002/cne.901990302

Van Essen, D. C., Newsome, W. T., & Maunsell, J. H. (1984). The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Research*, *24*(5), 429–448.

van Wezel, R. J. A., & Britten, K. H. (2002). Motion adaptation in area MT. *Journal of Neurophysiology*, *88*(6), 3469–3476. doi:10.1152/jn.00276.2002

Watamaniuk, S. N., & Sekuler, R. (1992). Temporal and spatial integration in dynamic random-dot stimuli. *Vision Research*, *32*(12), 2341–2347.

Watamaniuk, S. N., McKee, S. P., & Grzywacz, N. M. (1995). Detecting a trajectory embedded in random-direction motion noise. *Vision Research*, *35*(1), 65–77.

Watamaniuk, S. N., Sekuler, R., & Williams, D. W. (1989). Direction perception in complex dynamic displays: the integration of direction information. *Vision Research*, *29*(1), 47–59.

Watanabe, Y., Kezuka, T., Harasawa, K., Usui, M., Yaguchi, H., & Shioiri, S. (2008). A new method for assessing motion-in-depth perception in strabismic patients. *The British journal of ophthalmology*, *92*(1), 47–50. doi:10.1136/bjo.2007.117507

Watson, A., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A, Optics, image science, and vision*, *2*(2), 322–341.

Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., et al. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral cortex (New York, NY : 1991)*, *3*(2), 79–94.

Welch, L. (1989). The perception of moving plaids reveals two motion-processing stages. *Nature*, *337*(6209), 734–736. doi:10.1038/337734a0

Welchman, A. E., Lam, J. M., & Bülthoff, H. H. (2008). Bayesian motion estimation accounts for a surprising bias in 3D vision. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(33), 12087–12092. doi:10.1073/pnas.0804378105

Westheimer, G., & Truong, T. T. (1988). Target crowding in foveal and peripheral stereoacuity. *American journal of optometry and physiological optics*, *65*(5), 395–399.

Wheatstone, C. (1838). Contributions to the physiology of vision.--Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, *128*, 371–394.

Wilcox, L. M., & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research*, *49*(22), 2653–2665. doi:10.1016/j.visres.2009.06.004

Winawer, J., Huk, A. C., & Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychological science : a journal of the American Psychological Society / APS*, *19*(3), 276–283. doi:10.1111/j.1467-9280.2008.02080.x

Wright, M. J. (1987). Spatiotemporal properties of grating motion detection in the center and the periphery of the visual field. *Journal of the Optical Society of America A, Optics, image science, and vision*, *4*(8), 1627–1633.

Wright, M. J., & Gurney, K. N. (1992). Dependence of stereomotion on the orientation of spatial-frequency components. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, *12*(2), 264–268.

Xiao, D. K., Raiguel, S., Marcar, V., & Orban, G. A. (1997). The spatial distribution of the antagonistic surround of MT/V5 neurons. *Cerebral cortex (New York, NY : 1991)*, *7*(7), 662–677.

Xiao, D. K., Raiguel, S., Marcar, V., & Orban, G. A. (1998). Influence of stimulus speed upon the antagonistic surrounds of area MT/V5 neurons. *Neuroreport*, *9*(7), 1321–1326.

Yang, J., & Lisberger, S. G. (2009). Relationship between adapted neural population responses in MT and motion adaptation in speed and direction of smooth-pursuit eye movements. *Journal of Neurophysiology*, *101*(5), 2693–2707. doi:10.1152/jn.00061.2009

Yo, C., & Wilson, H. R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research*, *32*(1), 135–147.

Zeki, S. M. (1974). Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *The Journal of Physiology*, *242*(3), 827–841.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience*, *11*(3), 641–649.

Zhuang, J., Peltier, S., He, S., LaConte, S., & Hu, X. (2008). Mapping the connectivity with structural equation modeling in an fMRI study of shape-from-motion task. *NeuroImage*, *42*(2), 799–806. doi:10.1016/j.neuroimage.2008.05.036